DOCUMENT RESUME

ED 210 315 TM 810 972

| | |
|---|---|
| AUTHOR | Haney, Walt; And Others |
| TITLE | ESEA Title I Early Childhood Education: Review of Literature on Evaluation and Instrumentation. Interim Report, Revised. |
| INSTITUTION | Huron Inst., Cambridge, Mass. |
| SPONS AGENCY | Office of Education (DHEW), Washington, D.C. |
| PUB DATE | Jun 78 |
| NOTE | 170p. |
| EDRS PRICE | MF01/PC07 Plus Postage. |
| DESCRIPTORS | Compensatory Education; *Educational Assessment; *Educational Trends; *Federal Programs; Preschool Education; *Preschool Evaluation; *Program Evaluation; Program Improvement; *Testing Programs |
| IDENTIFIERS | Elementary Secondary Education Act Title I; Title I Evaluation and Reporting System |

ABSTRACT

Evaluations which describe and measure the impact of programs assisted under Title I are mandated by law. This evaluation is to be used in the decision-making process. A review of the literature on evaluation and instrumentation shows a paucity of descriptive information on early childhood programs funded under Title I. On-going evaluations of these programs are necessary. Services provided to pre-kindergarten, kindergarten, and first grade level students differ. Whole program efforts are offered at pre-kindergarten levels, while supplementary services are provided above and beyond children's regular school programs at the first grade level. Both whole program and supplementary program interventions are apparent at the kindergarten level. The differences in the structure of interventions funded at these levels, combined with the differences in program goals, illustrate the ineffectiveness of a single evaluation scheme for all three levels. High quality instruments for use in testing young children are scarce. Evaluation models and guidelines by which early childhood programs may be assessed are needed particulary in areas where no adequate instruments are available. (DWH)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED210315

SCOPE OF INTEREST NOTICE
The ERIC Facility has assigned this document for processing to: TM PS
In our judgement, this document is also of interest to the clearinghouses noted to the right. Indexing should reflect their special points of view. UD

Interim Report

# ESEA TITLE I EARLY CHILDHOOD EDUCATION: REVIEW OF LITERATURE ON EVALUATION AND INSTRUMENTATION

by

Walt Haney, et al.

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

TM 810 [illegible] 672

REVISED

INTERIM REPORT

# ESEA TITLE I EARLY CHILDHOOD EDUCATION: REVIEW OF LITERATURE ON EVALUATION AND INSTRUMENTATION

by

Walt Haney

With assistance of

Diane Schodlatz,

Deborah Walker,

&

Sara Weiss

June 1978

The Huron Institute
123 Mt. Auburn Street
Cambridge, Massachusetts
02138

The work reported here was performed pursuant to a contract with the USOE. Views expressed do not, of course, necessarily reflect those of the USOE or of the Huron Institute.

## PREFACE

This study has been carried out as part of Huron's project to investigate the feasibility of developing evaluation models for early childhood Title I programs. A draft version of the report was reviewed by several people who provided valuable suggestions: Robert Stonehill, USOE, Project Officer for this study; Deborah Walker of the Harvard School of Public Health; Peter Airasian of Boston College; Richard Zusman of the Massachusetts State Department of Education; Diane Proctor, of the National Testing Service Technical Assistance Center in North Carolina; Roy Hardy of the Educational Testing Service Technical Assistance Center in Atlanta,Georgia; and Sheldon White of Harvard University.

Suggestions of other project staff also contributed much to the improvement of the report: Mary Jane Yurchak, Co-Director of the Project; Anthony Bryk; and Richard Apling. Special thanks go to Susan Carbonell and to Marcia Tress, who typed much of the report and coordinated production of the rest of it.

TABLE OF CONTENTS


| | | PAGE |
|---|---|---|
| | Preface | i |
| I. | Background and Introduction | 1 |
| II. | Trends in Early Childhood Education | 6 |
| | 2.1 The Growth of Early Childhood Education in America | 8 |
| | 2.2 Parental Involvement | 15 |
| | 2.3 Comprehensive Services | 18 |
| | 2.4 State Coordination of Early Education | 21 |
| | 2.5 Individualization | 25 |
| | 2.6 Mainstreaming | 28 |
| | 2.7 Summing Up | 30 |
| III. | Evaluation of Federal Programs in Early Childhood Education | 32 |
| | 3.1 Broad Goals & Narrow Evaluations | 33 |
| | 3.2 Long-range Goals versus Short-range Evaluations | 36 |
| | 3.3 Program Definition and Implementation | 39 |
| | 3.4 Uncertain Findings | 42 |
| | 3.5 Rare Use of Evaluation Findings | 43 |
| | 3.6 Summing Up | 47 |
| IV. | Early Childhood Title I programs | 49 |
| | 4.1 Past National Evaluations of Title I | 50 |
| | 4.2 Scope of Early Childhood Title I Programs | 57 |
| | 4.3 Nature of Early Childhood Title I Programs | 62 |
| | • Prekindergarten<br>• Kindergarten<br>• First Grade | |
| | 4.4 Selection and Needs Assessment | 71 |
| | 4.5 Parental Involvement | 75 |
| | 4.6 Evaluations of Early Childhood Title I Programs | 78 |
| | 4.7 Summing Up | 86 |

PAGE

V. Early Childhood Tests and Instruments 89

5.1 Instruments Used in Evaluations of Title I Early Childhood Programs 89

5.2 Characteristics of Young Children that Make Testing Difficult 94

5.3 Types of Instruments 95

5.4 Technical Quality 108

5.5 Test Use 120

VI. Summary and Conclusions 127

Appendix I: State Title I Evaluation Reports for 1975-1976 by State for 1975-1976 137

Appendix II: General Bibliography 147

## TABLES & FIGURES


| | | Page |
|---|---|---|
| Figure 1.1: | Total Children 3-5 Years Old and Number Enrolled in Preprimary Programs: United States 1964-1975 | 11 |
| Figure 1.2: | Average Number of Articles Per Year | 14 |
| Figure 1.3: | Average Number of Articles Per Year - Parent Education | 16 |
| Figure 1.4: | Average Number of Articles Per Year - Individual(ized) Instruction | 26 |
| Table: 4.1 | Attention Given to EC Programs in National Evaluations | 51 |
| 4.2 | Most Convincing Successful Programs | 54 |
| 4.3 | Title I Participants by State PK-K and PK-12 | 59 |
| 4.4 | Summary of Review of FY 1976 State Title I Evaluation Reports: Numbers of Participants | 60 |
| 4.5 | Percent of Title I Participants in PK, K, and Grade 1 | 67 |
| 4.6 | Iowa Criteria | 72 |
| 4.7 | Summary of Review of FY 1976 State Title I Eval. Reports | 83 |
| Table: 5.1 | Tests Reported Used in Title I Evaluations of EC Levels | 90 |
| 5.2 | Frequency with Which Standardized Tests Cited in General Review | 92 |
| 5.3 | Major Sources of Information in Tests | 96 |
| 5.4 | Early Childhood Tests and Instruments Listed in Buros' Seventh Mental Measurements Yearbood (1972) | 98 |
| 5.5 | Preschool and Kindergarten Subtests | 100 |
| 5.6 | Subtests for Grades 1-4 | 101 |
| 5.7 | Unpublished Early Childhood Instruments (1971) | 103 |
| 5.8 | Unpublished Early Childhood Instruments (1976) | 104 |
| 5.9 | Socioemotional Measures for Preschool and Kindergarten Children | 107 |
| 5.10 | MEAN Evaluations of Preschool and Kindergarten Subtests | 115 |
| 5.11 | MEAN Evaluations of Subtests for Grades 1-4 | 116 |
| 5.12 | Classification of Selected Subtests as Cognitive or Readiness Instruments in CSE-ECRC Test Evaluations | 123 |

## I. BACKGROUND AND INTRODUCTION

In the last decade, the United States Office of Education has experienced frequent frustration in its attempts to evaluate systematically the effectiveness of programs funded under Title I of the Elementary and Secondary Education Act of 1965 (ESEA). As a result, in the Education Amendments of 1974, the U.S. Congress authorized the Commissioner of Education to:

> "(a)...provide for independent evaluations which describe and measure the impact of programs and projects assisted under this title. Such evaluations may be provided by contract or other arrangements, and all such evaluations shall be made by competent and independent persons, and shall include, whenever possible, opinions obtained from program or project participants about the strengths and weaknesses of such programs or projects."
>
> "(b)...develop and publish standards for evaluation of program or project effectiveness in achieving the objectives of this title."
>
> "(c)...consult with State agencies in order to provide for jointly sponsored objective evaluation studies of programs and projects assisted under this title within a State."
>
> "(d)...provide to State educational agencies, models for evaluations of all programs conducted under this title... which shall include uniform procedures and criteria to be utilized by local educational agencies, as well as by the State agency in the evaluation of such programs."
>
> "(e)...provide such technical and other assistance as may be necessary to State educational agencies to enable them to assist local educational agencies in the development and application of a systematic evaluation of programs in accordance with the models developed by the Commissioner."
>
> "(f)... specify objective criteria which shall be utilized in the evaluation of all programs and shall outline techniques (such as longitudinal studies of children involved in such programs) and methodology (such as the use of tests which yield comparable results) for producing data which are comparable on a statewide and nationwide basis."
>
> (P.L. 93-380, Section 151).

In accordance with this mandate, the Office of Education sponsored the development of a set of evaluation models as a basis for local and state reporting of comparable Title I evaluation data. Initial versions of these Title I evaluation models were developed by the RMC Research Corporation in 1975, discussed with state education agencies in 1976, and further refined in briefing sessions in 1977. Efforts are currently underway at USOE to draft regulations describing intended use of the standard evaluation models or approved alternatives. The three basic types of evaluation models elaborated by RMC investigators are:

A. Norm-Referenced Design

B. Control Group Design

C. Special Regression Design

These three types of models - each having variations allowing the use of normed or non-normed tests - are described in the USOE publication, User's Guide ESEA: Title I Evaluation and Reporting System (Tallmadge and Wood, 1976). Under this system, results of evaluations at the local level are to be aggregated across local programs, and eventually across states, by converting estimates of project impact to a common metric - normal curve equivalents or NCEs. This equating is to be accomplished by comparing test results to national norms provided by standardized tests of reading and mathematics.

For several reasons, this evaluation and reporting system is not usually applicable to Title I programs serving children below the second grade. First, few tests for children prior to the end of first grade have adequate national norms. Second, the models are designed to assess program impact on childrens' achievement in reading, mathematics and language arts. While these domains

encompass the objectives common to most Title I programs for grades 2-12, Title I programs for very young children are much more likely to have more diverse goals - including reading readiness, social behavior, and psychomotor development. Third, the OE system for evaluating Title I grades 2-12 programs is aimed primarily at annual cross-sectional evaluation, but there is much concern of late for longitudinal evaluations of educational interventions with young children (Lazar, 1977).

Because the previously developed evaluation models are not applicable to Title I programs serving children below the second grade, and because of requests from local and state education agencies for guidance on early childhood evaluation, the USOE decided to fund a project to develop evaluation procedures for early childhood programs. In the spring of 1977, the USOE requested proposals for a project to consider the development of evaluation models for Title I programs in early childhood education, (USOE, RFP 77-79). The Huron Institute was awarded a contract for this project in September 1977. This literature review comprises the first phase of the project.

The purpose of the literature review is to provide background information for Huron's investigation of the feasibility of developing models for evaluating early childhood education programs funded under Title I. In preparing this review, we have surveyed four domains of literature:

1) Early childhood education literature in general;
2) Evaluations of Federal programs concerned with early childhood education;
3) Evaluations of Title I early childhood education programs; and,
4) The literature concerning tests and instruments for young children.

The next four sections of this literature review correspond to these four areas of review. The final section of the review summarizes

what we have learned and attempts to identify implications of our findings from our further work in conducting the feasibility study.

Before launching into the substance of our literature review, we must pause to recount a problem which has troubled many of our efforts - namely the ambiguous meaning of several words and phrases with which we must deal. The project of which this literature review is a part is entitled "Development of Evaluation Models for Title I Programs in Early Childhood Education". The term 'early childhood education' is used here to refer to educational endeavors in the prekindergarten through first-grade range. Yet according to the National Center for Educational Statistics' Combined Glossary, the term is more generally defined as follows:

> Early Childhood Education: Educational experiences provided by the school at the preprimary and primary levels (NCES, 1974, p.43).

Hence, the term, 'early childhood education', or ECE is used in this paper in two different ways. As used with regard to this particular project, particularly in chapters IV, V, and VI, the phrase refers to the prekindergarten, kindergarten and first grade grade-levels, roughly corresponding to ages three-to-four, five and six. However, when the term is used with reference to the education literature in general in chapters II and III, it implies a broader grade-and-age-span, roughly prekindergarten to third grade, and ages three through eight.

Another term carrying considerable ambiguity is "preschool". In some cases, "preschool" is used to refer to both prekindergarten and kindergarten. In others, it seems to have been used to designate only the prekindergarten level, and in still other cases, especially in

states without state-wide kindergarten programs, the term seems to have been used to designate pre-first-grade programs for five-year-olds, or what would elsewhere be called kindergarten. In chapters II and III, when dealing with the early childhood literature generally we will simply have to tolerate a certain amount of ambiguity in use of the terms "early childhood education" and "preschool". In chapters IV - VI, however, we use the term ECE to refer strictly to the prekindergarten to first grade grade levels and attempt to avoid the term preschool and rely, as much as possible, on the differentiated terms "prekindergarten", "kindergarten", and "first grade". As we progress through the Literature Review, we will, where necessary, further define the terms used.

## II. TRENDS IN EARLY CHILDHOOD EDUCATION

The funding of early childhood programs under Title I is a part of a broad upsurge of interest in early childhood education in America. The purpose of this chapter is to describe this broader movement; specifically to document the growth of interest in early childhood education and to identify several broad trends apparent in the early education movement. The effort is premised on the simple notion that as early childhood Title I programs have been affected by broader currents of early childhood education in America in the past, so they will be in the future. Hence, in considering how early childhood Title I programs might be evaluated in the future, one should not restrict attention to just what these programs are now, but ought also to consider what they may be in the future. Trends apparent in early childhood education in general may provide clues to that future.

Before turning to the substance of the chapter, let us offer several comments on the thinking behind our approach and several cautions on that thinking. The identification and documentation of societal trends is no easy matter. We may have no inkling today of what may seem patently obvious tomorrow. Even if we can agree on what we see as broad currents in educational affairs, documenting their existence is not always an easy task. Hence, in this chapter we have adopted two approaches to identifying trends in early childhood education - appeals to expert opinion and appeals to numbers. In the first approach, we have simply surveyed recent early childhood education literature to identify trends which experts in early education perceive. In addition to this appeal to authority, we have tried, where possible, to back expert opinion up with facts and figures.

The information from both these sources, that is expert opinion and available data, has an uncertain relationship to what is actually happening in schools and communities across the nation. Therefore, it is

important to keep in mind that the trends we can identify from the literature may not correspond very closely to trends in practice. In short, there may be a large gap between educational literature and educational practice, and whatever the relationship between literature and practice is, it likely is not stable over time. Nevertheless, the trends identified below may foreshadow something of the future of early childhood education programs.

The chapter is organized into seven sections. The first simply documents the upsurge of interest in early childhood education in the last twenty years. The second recounts the special prominence given to parental involvement in early childhood programs. The third documents the trend toward comprehensive approaches in early programs. The fourth reports a possible trend toward increasing state coordination of early education. The fifth documents the trend toward individualization in early programs and the sixth section discusses the recent enthusiasm for mainstreaming of special needs children. A final section sums up what's been said and suggests some of the implications that these trends may hold for evaluating early childhood Title I programs.

## 2.1 The Growth of Early Childhood Education in America

Early childhood education in America has a long history. The roots of recent concern for early education can be traced back to colonial times to the resurgence of concern for public education in the 19th century (Braun and Edwards, 1972), and to the growth of nursery school instruction in the 1920's (Kamii, 1971). The federal government sponsored nursery schools in the 1930's as a means of creating work for unemployed teachers, and during World War II the government funded day care centers for children of working mothers employed in war industries. The latest upsurge of interest in early education dates back, however, only to the 1960's (Kamii, 1971).

Two factors have been especially influential in promoting concern for early education since that time--research findings and governmental support. Research findings seem to have provided the initial catalyst for increased efforts in early education. Urie Bronfenbrenner, writing in 1974, maintained that "early intervention... as a strategy for counteracting the destructive effects of poverty on human development... had its roots in an emergent body of theory and research in the 1950's pointing to the beneficial effects of early stimulation both in animals and humans." (Bronfenbrenner, 1974, p.1). The belief that early childhood is a critical period for intervention was embodied in frequent discussions of the plasticity of early development (Elkind and Lyke, 1975; Evans, 1975; Weber, 1970). This term--the plasticity of early development--was used to refer to the idea that in their early years, children's development is more malleable, and hence more likely influenced by intervention than in later years. The work of three psychologists - Piaget, Hunt, and Bloom - was especially influential in the 1960's in promoting this idea (Braun and Edwards, 1972; Center for Educational Research and Innovation, 1975; Evans, 1975; Lee, 1973; Weber, 1973).

Piaget's work, which became widely known in this country in the 1960's, directed attention to the fundamental processes governing changes in children's early development. Piaget's theory of cognitive stages suggested that children's stage development is fastest in the early years - they are thought generally to pass through three of four qualitatively different cognitive stages in the first 12 to 14 years of life (Collis, 1975). Hunt (1961), in his influential book Intelligence and Experience, argued that children's intelligence is significantly affected by the kinds of encounters they have with their early environment. Finally, Bloom's (1964) summary and analysis of longitudinal data concluded that the rate of intellectual development is greatest during early childhood. On the basis of his longitudinal analysis of IQ scores, he contended that "about 50% of intellectual development takes place between conception and age 4". (Bloom, 1968, p. 88). The idea of plasticity of early development prompted an unprecedented concern for children's early cognitive development. Early childhood, went the thinking, is a critical period in which to intervene in order to maximize children's intellectual development to full potential. This thinking provided much of the rationale for early compensatory education.

The research on child development and cognition coincided with a second major influence - the spirit of social reform that pervaded the politics of the early 60's. Early childhood education became a focus of the social reform movement of the 1960's--what is commonly called the War on Poverty. In 1965, Project Head Start was funded to provide pre-school experience for poor or disadvantaged children and quickly became the showpiece of the War on Poverty. The objectives of Head Start were broad, including provision for food and health services, preschool education and parent education (Grotberg, 1969).

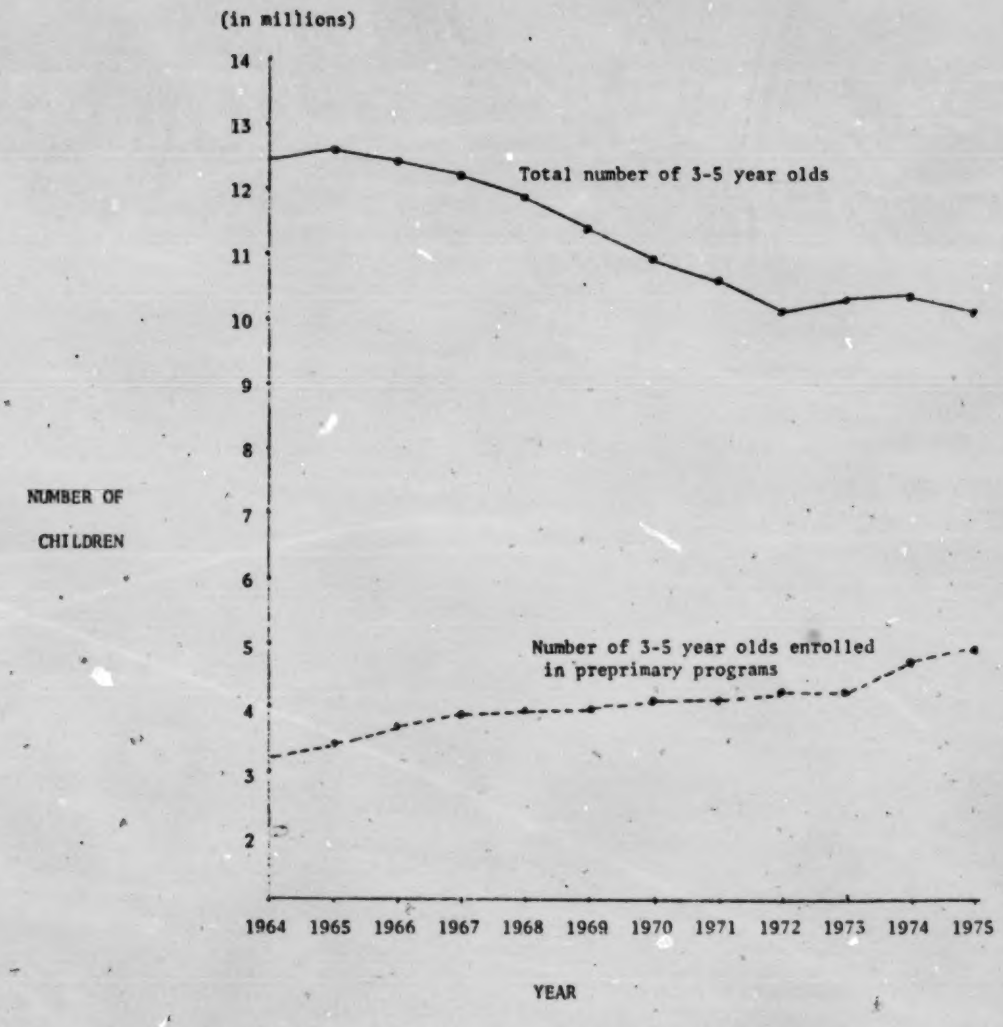
The hope was that Head Start and other early interventions would permit disadvantaged children to "catch-up" by providing what home environments did not provide--good nutrition, health care, intellectual stimulation and socialization. The growing national commitment to early childhood education for the poor was reflected in the amount of Federal money invested in preschool programs. From 1960 to 1972, federal investment in preschool programs increased from near zero to approximately $1 billion (Stocker, 1973). At the same time, the number of middle class nursery schools grew and more states began to offer public kindergarten programs (Elkind and Lyke, 1975; Lazerson, 1972).

Two simple ways to document the increasing activity in early childhood education are to note 1) the number of children enrolled in early childhood programs in the past decade and 2) the number of professional articles published in that time concerning young children. Figure 1.1 shows the preprimary enrollments compared to total population of children aged 3 to 5 years from 1964 to 1974.* The enrollment figures include both private and public programs. Although the total number of the 3-to-5-year-olds steadily declined in this decade, the enrollment in preprimary programs has steadily increased. This has resulted in an increase in the proportion of children aged 3 to 5 years attending preprimary programs from 30 percent in 1964 to nearly 50 percent in 1975. (Barker and Hurd, 1971; Grant and Lind, 1976).

---

*Preprimary enrollment figures were not collected by the federal government until 1964, a fact which itself reflects the upsurge of interest in early childhood education in the sixties.

Figure 1.1: Total Children 3-5 years old and number enrolled in pre-primary Programs: United States, 1964-1975.




Source: Barker & Hurd, 1971; Grant & Lind, 1976

Most of this increase, it should be noted, stemmed from increases in enrollment rates at the prekindergarten level. Enrollment at the pre-kindergarten or nursery school level more than doubled from 1967 to 1977, while enrollment at the kindergarten level remained relatively stable. In fact, in 1977 total kindergarten enrollment declined for the first time in 30 years. According to the Bureau of the Census, the eight percent decline in kindergarten enrollment from 1976 to 1977, parallels a drop in the number of births five years ago. (U.S. Bureau of the Census, 1978). Since the kindergarten enrollment rate is now more than 90%, the Census Bureau predicts that future changes in kindergarten enrollments will be influenced more by birth rates than by shifts in enrollment rates. However, since the pre-kindergarten enrollment rate is still far less than 50%, it seems likely that changes in numbers of children attending prekindergarten will continue to be affected both by birth rate and enrollment rate changes.

Figure 1.2 shows the number of articles listed in the Education Index under the headings of "kindergarten", "nursery schools", and "preschool education" from 1930 to 1977* The number of articles concerning kinder-garten has remained between 25 and 40 per year except for a slump during

---

*The average annual number of entries in the Education Index is, of course an imperfect indicator of educators' interest in any particular topic. Therefore, a couple of caveats regarding the adequacy of this measure are in order. First, we might ask whether changes in the number of articles in educational journals on any given topic reflect changes in relative interest in the topic or whether it only mirrors changes in the number of educational journal articles on all topics. While the latter supposition may be true to some extent, growth in the total number of educational journal articles clearly does not account for all of the changes reflected in Figure 2. The size of the Education Index grew from about 630 pages per year in 1930 to slightly over 1000 in recent years (with a roughly constant number of entries per page). Also, note that the Index was published triannually from its founding in 1930, to 1953, biannually until volume 13 in 1963, and annually thereafter. Another problem with using entries in the Index as a measure of interest among educators is the inconstancy of the indexing system. In the text below we elaborate on these caveats.

the war years, whereas the number of articles concerning nursery schools has declined steadily from a high of 31 in 1931 to zero in 1977. In part, this decline could reflect the increasing categorization of articles on nursery-school-age children under the broader term preschool education. The number of articles categorized under the latter term increased dramatically around 1964--coinciding with the inception of Project Head Start. From 1930 to 1960 the annual number of articles on preschool education averaged only five to ten. By 1964, however, the figure had zoomed all the way to 86, and since then more than 50 articles per year on preschool education have been listed in the Index.

Thus, both enrollment figures and the education literature show a dramatic increase in concern for early education over the past two decades. Amidst the increased interest in early childhood education in general, several themes have received increasing attention. In the sections that follow, we document five trends that seem to have had special prominence and seem also to have special relevance for the future of Title I early childhood programs and for future evaluations of such programs.

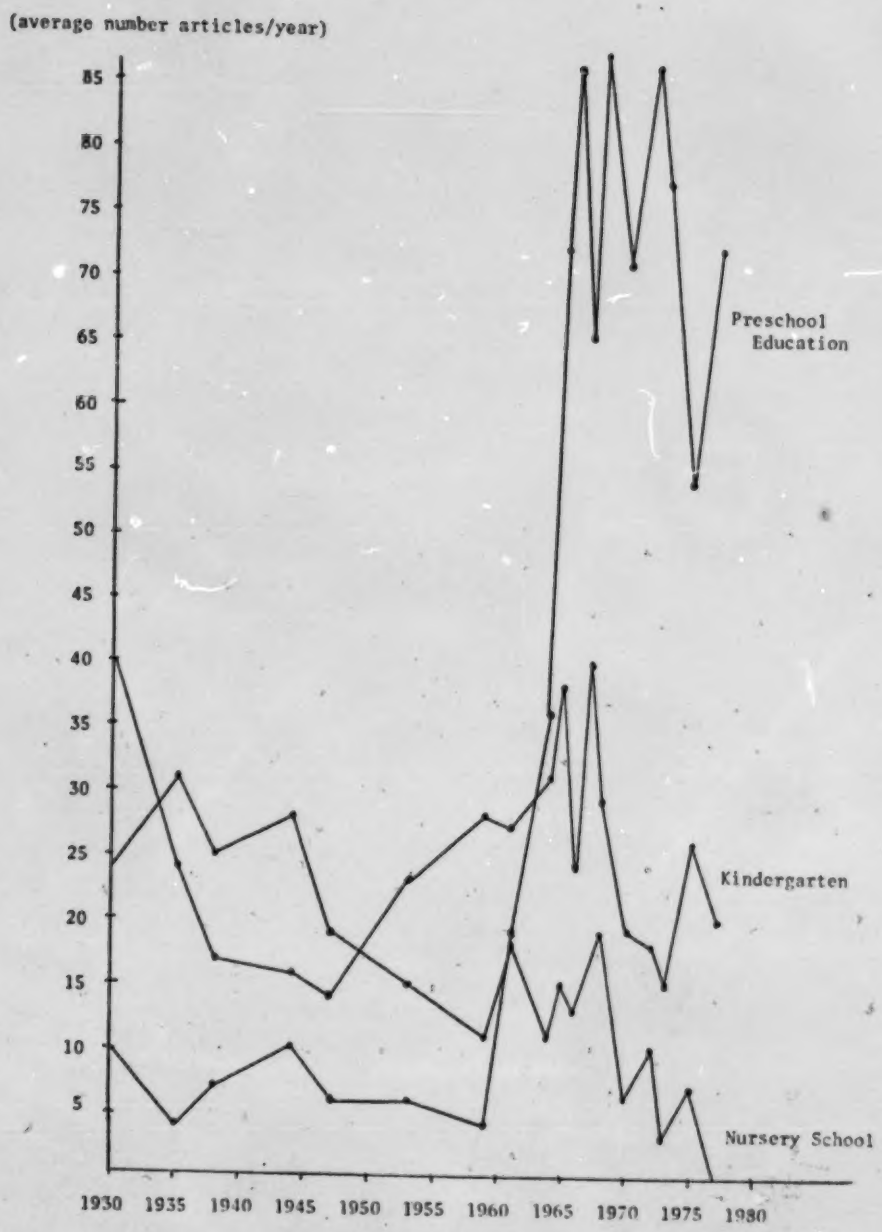



Figure 1.2: Average number of articles/year listed under "Kindergarten", "Nursery School", and "Preschool Education" in The Education Index at specified dates between 1930 and 1977.

## 2.2 Parental Involvement

The clearest trend in early childhood education since the early 1960's is a continuing emphasis on parental involvement in programs for young children. The importance of involving parents in programs of early intervention seems to be a prominent concern of every major writer on early childhood education (for example, Bronfenbrenner, 1974 ; Hunt, 1975). The idea that parents should be involved in efforts to spur the development of young children is expressed in a variety of terms. Parent involvement, parent training, parent participation, and parent education are a few of the most common.

The increasing emphasis on parent involvement is evident in the education literature. Figure 1.3 shows the annual number of articles categorized under the heading "parent education" in the Education Index from 1931 to 1977. Parent education appears to have been a topic of concern in 1930, with interest decreasing somewhat in the 30's and remaining at a relatively low level in the 40's, 50's and 60's. Then in the 1970's the number of articles concerning parent education quadrupled.

The terms parent involvement and parent education encompass a variety of concerns. McLaughlin (n.d.), for example, distinguishes between "parent participation" in school-sponsored activities or on advisory councils and "parent training", that is, training of parents in child development, early education and other areas. Hess (1973) suggests that the manner in which parents are involved may vary according to the ages of their children, with the emphasis on training for parents of younger children, and on participation for parents of children closer to school age. Gordon, who has written extensively on parent education suggests that the extent of parent involvement conditions the manner

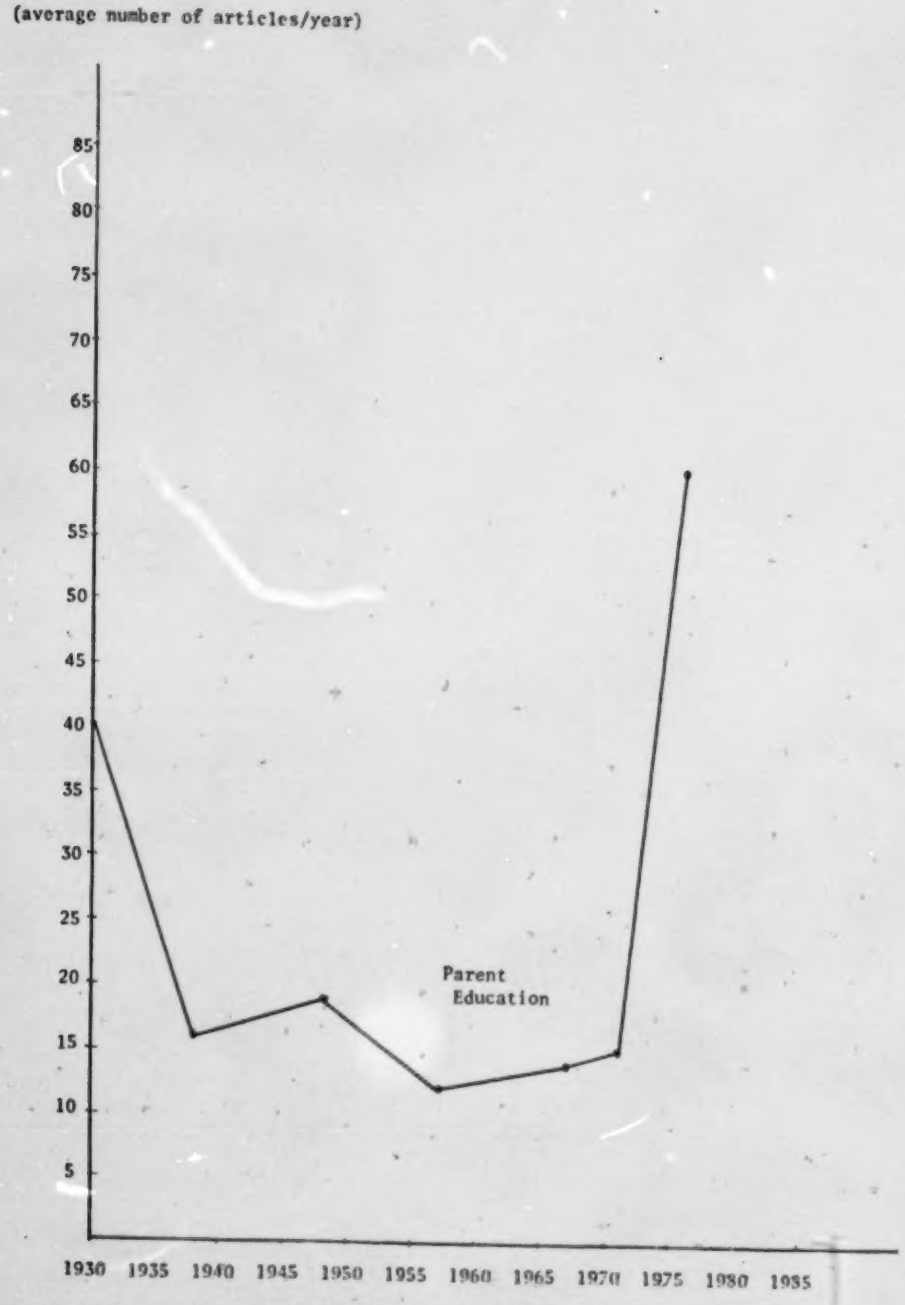



Figure 1.3: Average number of articles/year listed under "Parent Education" in The Education Index of specific dates between 1930 and 1977.

of that involvement. Center-based programs for 4-and-5-year-old children often use parents as volunteers or paid paraprofessional staff to actively teach their children in a classroom setting. In center-based programs, parents can also be involved in decision-making at all levels of planning and can serve as resources for professional educators. In home-based programs, emphasis seems likely to be more on parent training than on participation.

Federal education programs have sought to spur both parent participation and parent training. Parents have both served as classroom aides and participated in advisory councils for Projects Head Start and Follow Through. Parent-Child Centers and Project Home Start, funded by the Office of Child Development, seek to spur parental involvement in terms of participation and training.

Lois-Ellen Datta (1975) suggests that the trend toward increased concern with parent involvement is based on six factors: 1) preliminary evidence that parents determine more of their children's educational achievement than does the school; 2) increased concern for the first five years of life when parents traditionally have responsibility for a child's development, especially concern that early intervention should occur before the ages of 4 or 5 years - even before 18 months (White, 1973) - if it is to have the desired, long-term effects; 3) failure of school-based systems to deliver equality of educational outcomes to disadvantaged children (Jencks et al., 1972); 4) increased demand for citizen participation in decision-making in all sections of society (Hess et al., 1971); 5) increased appreciation of what schools can learn from parents, and an emphasis on the positive aspects of cultural diversity; and (6) increasing awareness of the need for parent education by the general public.

Evidence in support of parent involvement in early education has been summarized by Urie Bonfenbrenner. After reviewing several dozen longitudinal studies of early intervention, Bronfenbrenner put forth the following as a general principle of early intervention.

> The evidence indicates that the family is the most effective and economical system for fostering and sustaining the development of the child. The evidence indicates further that the involvement of the child's family as an active participant is critical to the success of any intervention program. Without such family involvement, any effects of intervention, at least in the cognitive sphere, appear to erode fairly rapidly once the program ends. In contrast, the involvement of the parents as partners in the enterprise provides an on-going system which can reinforce the effects of the program while it is in operation, and help to sustain them after the program ends. (Bronfenbrenner, 1974, p. 55).

Parent involvement in early education programs has not, however, been fully endorsed by everyone. For example, in a recent article in the Teachers College Record, Schlossman (1978) maintains that the increasing emphasis in recent federal programs on parent education as opposed to parent participation, "shift(s) the burden of accountability for failure from the government-sponsored professional education to the poverty parent". Arguing that the recent federal enthusiasm for parent education embodies a "remarkably anachronistic view, namely that parents wholly determine their children's futures", Schlossman warns that "It is fruitless to try to isolate the impact of individual educational institutions--whether the family, the school, or the television--and assign them full responsibility for educational outcomes".

2.3 Comprehensive Services

A second major theme in the early education literature, directly related to Schlossman's criticism, concerns the need for comprehensive services to meet the full needs of disadvantaged

children. The term comprehensive services refers to the use of multi-disciplinary approaches to meeting the needs of children, the coordination of service provision and cooperation among service agencies (for example, those providing mental health, educational health, social, nutritional and economic services in a community).

In part the concern for comprehensive services has grown out of the history of early childhood education in America, especially the tradition often characterized as the "whole-child approach." This approach holds that early childhood education must attend to all aspects of children's development--cognitive, socioemotional, physical, and psychomoter. The whole child approach has characterized traditional, middle class nursery schools since their inception in the 1920's and clearly was influential in the founding of Project Head Start (Branch and Overly, 1971; Evans, 1975). One example of the whole-child approach to early childhood education is the development interaction viewpoint on early education that has evolved over the past fifty years at the Bank Street College of Education in New York City (Biber, Shapiro, and Wickens, 1971). This viewpoint combines both traditional and contemporary thinking about child development and early childhood education. It emphasizes the interaction of cognitive and affective development, specifies the requirements of a good learning environment and relies heavily on direct observations for use in formative evaluations. The Bank Street tradition tends not to differentiate compensatory education or early intervention from early childhood education in general, but instead holds out the same educational goals apply to all children.

At present, the whole-child approach to early childhood education has wide support (Barnes, 1976; Bauch and Shearron, 1969; Gallagher, 1974; Harty and Roberts, 1973; Leeper, 1976; McIntyre, 1975; Osborn, 1975; Ramptom, 1973; Riles, 1973; Stearns, 1970; Weber, 1970).

Concern for meeting the full needs of children with a range of services has been especially prominent in compensatory education efforts because of the special needs of disadvantaged children. Such children and their families lack many basic resources, health care and good nutrition being two primary ones. From the beginning, the objectives of Project Head Start were to improve not just children's school performance but also their physical health and well-being. Hence Head Start has provided both medical and nutritional as well as educational services. Early evaluations of Head Start such as the Westinghouse Study (Cicirelli, et al., 1969) neglected those benefits but more recent reviews, for example White et al, (1973), have continued to emphasize the relationships between physical health, nutrition, mental health and learning--in short to reaffirm the need for comprehensive services to disadvantaged children. Recently concern for comprehensive services to disadvantaged children has been expressed in terms of the need for ecological approaches to early intervention. For example, Bronfenbrenner in 1974 wrote:

> The research results indicate that for the children from the most deprived groups no strategy of intervention is likely to be effective that focuses attention solely on the child or on the parent-child relationship. The critical forces of destruction lie neither within the child nor within his family but in the desperate circumstances in which the family is forced to live. What is called for is intervention at the ecological level,measures that will effect radical changes in the immediate environment of the family and the child. (p. 54)

Concern for comprehensive services in terms of an ecological approach to early intervention is reflected in the renaming of the Office of Child Development as the Administration for Children Youth and Families. Also, in 1976 the federal government's Interagency Panel on Early Childhood Research and Development selected ecological studies of child development as a theme to guide its coordination efforts.

Ecological approaches to providing comprehensive services may be an on-going theme in early childhood education, but the future of this trend is not altogether clear. The "back-to-basics" sentiment of the late 1970's has focused mainly on improving reading, writing and motor skills of secondary school students, but there is some evidence that this conservative educational trend may have impact on early education as well. Sentiment along these lines was expressed, for example, in a recent report by a panel of the National Academy of Education:

> The Panel is in agreement that a series of standardized tests at the lower grade levels used for diagnosing individual student weaknesses, pinpointing remediation needs, and building public pressures if school-wide performances in basic skills continue over time to be consistently low, could be positive influences on student learning. (NAE, 1978, p.v.)

It is unclear how such concern over basic skills achievement in the late 1970's will affect the longer-term trend toward comprehensive services in early education. It may be, for example, that concern for comprehensive services will be maintained but perhaps with priority given to those services which seem to afford most potential for boosting basic skills achievement.

## 2.4 State Coordination of Early Education

Related to the concern for coordinated services is the trend towards making early childhood education public and coordinating educational programs

at the state level. Much work has already been done to coordinate services to children in many states. The Education Commission of the States devoted an entire issue of its periodical Compact in August 1973 to reporting on the progress being made by various states such as Massachusetts (Liedermann, 1973), Utah (Rampton, 1973), and California (Riles, 1973) in statewide coordination of early childhood education.

As of mid-1973, 14 states* had established child development offices and four** more had placed centralized authority for child development activities in their state education departments. Evans (1975) refers to this as the "task force phenomenon" in which specialists in child development and education have gathered to recommend master plans for state-wide early childhood education. The most well-known of these task forces is that of the State of California (1972), which recommended public early childhood education beginning at age four (voluntary) and extending until age eight. The task-force plan for ECE in California featured: 1) local development of programs; 2) parent participation; 3) early identification, diagnosis and remediation of learning disabilities; 4) individualized instruction; and, 5) the use of primary schools as community centers for the coordination of physical, social and psychological services. State legislation to implement this plan passed in 1972.

---

*Arkansas, Florida, Hawaii, Idaho, Louisiana, Massachusetts, Mississippi, North Carolina, South Carolina, Tennessee, Texas, Vermont, Washington, West Virginia.

**Alaska, California, Georgia and Utah.

Of all early childhood programs available, public kindergartens seem to have the widest support (Elkind, 1974; Sigler, Keyserling and Hedges, in Stocker, 1973), and it is projected that by 1980 all fifty states will have public kindergartens. In 1974, 79 percent of the five-year-old population already were enrolled in public or private kindergartens (King, 1974) compared to only 69 percent just five years earlier (Barker and Hurd, 1970). A number of organizations - the Council of Chief State School Officers, American Association for School Administrators, American Federation of Teachers, National Academy for School Executives - have backed public preschools for four-year-olds, while the latter two even support public education for three-year-olds (Stocker, 1973). Total preprimary enrollment of children ages three to five is projected to increase from 4.5 million in 1974 (King, 1975) to 7.6 million in 1980, with 100 percent of all five-year-olds, 50 percent of all four-year-olds, and 30 percent of all three-year olds being enrolled (Stocker, 1973).

The future of this trend toward state coordination of early childhood education is not altogether clear. Considerable concern has been expressed over the form that governmental coordination of early education should take. For example, psychologists and early childhood educators like Zigler (1973) and Weber (1970) have expressed fear that public school control of preschool programs will result in a downward extension of negative aspects of primary schooling. These observers fear that an excessive focus in early childhood education on academic readiness and cognition, may lead to inattention to children's socio-emotional development. On the other hand, primary educators and public school personnel (Barnes, 1976; Bauch and Shearron, 1969; Frazier, 1968; Harty and Roberts, 1973; Stocker, 1973) feel more positive

about the public primary schooling and think that primary education itself is changing in directions that make it more congruent with early childhood education. For example, Barnes (1976), Bauch and Shearron (1973) and Harty and Roberts (1973) all report trends towards individualized instruction, affective education and process or method-oriented teaching (as opposed to content- or subject-matter oriented teaching) in the elementary schools. Regardless of their orientation for or against public school control of early childhood education, educators of assorted stripes have voiced concern for some sort of coordination between early childhood education and elementary schooling (Colman and Sandoval, 1971; Frazier, 1968; McIntyre, 1975; Weber, 1970).

With the growing concern over basic skills achievement in elementary schools mentioned above, this issue of coordination seems likely to remain prominent.

## 2.5 Individualization

A fourth major theme in early childhood education, individualization, reflects a general trend in education that seems to have increasing and relatively unclouded support. Individualization, or individualized instruction refers to the development of flexible educational programs designed to fit individual children's needs instead of across-the-board curricula applied alike to supposedly homogeneous groups of children, such as the disadvantaged. Individualization emphasizes the unique characteristics of different children, each with his or her own pattern of strengths and weaknesses.

The increased emphasis on individualization is clearly reflected in the education literature. Until 1969, articles on individualized instruction were listed under "individual instruction" in the Education Index and subcategorized under "kindergarten-elementary schools" and "secondary schools". In 1969, the main heading was changed to "individualized instruction". Figure 1.4 shows the number of articles listed under "individual(ized) instruction/kindergarten - elementary education" from 1930 to 1976. As reflected in these data, interest in individual instruction gradually increased from the 1940's until the 1960's. Then, in the last decade, interest in individualization escalated sharply. From 1967 to 1975 articles on this topic increased more than three-fold from 20 to roughly 60 annual listings in the Education Index. The earlier, gradual growth of interest corresponds to the period when elementary educators began to group children by ability and to instruct these groups separately (Lee, 1973). In the 1960's and thereafter, individual(ized) instruction seems to have focused more on programming for the unique needs of individual children rather than ability grouping.

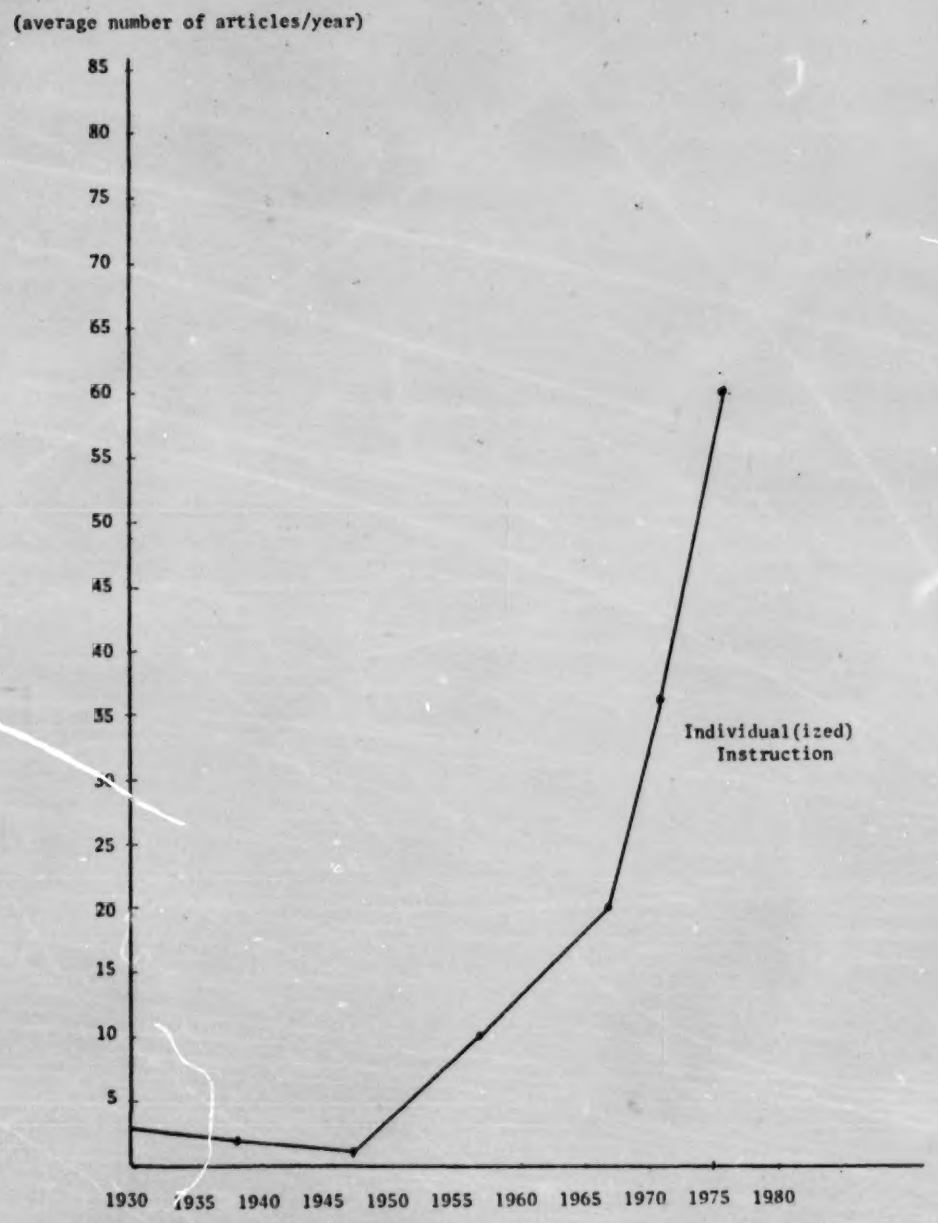



Figure 1.4: Average number of articles per year listed under "Individual(ized) Instruction" in The Education Index between 1930 and 1977.

Related to the notion of individualization is the concept of diagnostic-prescriptive teaching, which represents a development especially prominent in special education (Gallagher, 1974). Diagnostic-prescriptive teaching refers to a continuing process of evaluation of the needs of individual children, and teaching directed towards those needs. In early childhood special education, the goal of diagnostic-prescriptive teaching usually is the prevention of problems and the remediation of already existing problems before they become more serious. The means to prevention in this approach are early screening, early identification and diagnosis, and prescription of appropriate remediation. Numerous screening instruments have been developed during the past decade in efforts to identify children who are likely to experience difficulty in school before they encounter formal academic training (Mardell and Goldenberg, 1972; Northeast Regional Research Center, 1975; Nuttall and Gomes, 1975). More than ten of the fifty states have plans for statewide health and developmental screening for children in the first five years of life (Sulzbacher and Kenowitz, 1977), and an even larger number of states have mandated prekindergarten (ages 4 1/2 to 5 1/2 years) screening for children with potential learning problems (Stone, 1976). Along with individualization, early screening is a definite trend in early childhood education.

## 2.6 Mainstreaming

Another changing aspect of early childhood education originated with older children in special education programs.* This is the concept of mainstreaming or normalization (Hobbs, 1975; Wolfensberger, 1972). The concept refers to the integration of children with special needs into regular classrooms instead of segregating and thereby stigmatizing them in special classes. The push for mainstreaming began in the courts and in the state legislatures, rather than in the schools. Proponents of mainstreaming argue that special needs children in the past have often been improperly isolated in special classes, thereby preventing them from receiving equal educational opportunities (Divoky, Snyder, Appoloni and Cooke, 1977).**

Though mainstreaming began as a special education issue, it has, in recent years, had impact far beyond the bounds of special education. On the national level, Project Head Start regulations were changed in 1972 to require all Head Start centers to include children with special needs as not less than 10 percent of their total population. More recently, the Education for All Handicapped Children Act of 1975, Public Law 94-142, requires the states to establish procedures to assure "that to the maximum extent appropriate, handicapped children...are educated with children who are not handicapped" (Section 612 (5)B).

Mainstreaming has relevance to early childhood Title I programs for a variety of reasons. First, mainstreaming is viewed as expecially relevant

---

*For example, PARC v. the Commonwealth of Pennsylvania and Mills v. Board of Education of the District of Columbia; The Special Education Act of 1972 of the Commonwealth of Massachusetts, Chapter 766.

**The mainstreaming movement in special education actually has historical and intellectual ties to the racial desegregation movement. We do not have space here, however to elaborate on this connection.

for young children, as a means for preventing negative attitudes toward the handicapped or disadvantaged from developing in early life. Also, it is viewed as a way to avoid the trauma of mainstreaming later in life, after a period of segregation, and to prepare children and parents for mainstreaming in the elementary and secondary schools (Divorky; 1976). More directly, the issue of mainstreaming has been brought to the fore of the debate over the future of TitleI. In a recent paper, Gene Glass and Mary Lee Smith (1977) have argued strongly against pull-out procedures whereby children are withdrawn from regular classes in order to receive remedial instruction. In the past, pull-out procedures have been widely used in Title I programs. However, Glass and Smith, using precisely the same sort of logic as advocates of mainstreaming in special education, argue that "Instructional strategies should be devised that would eliminate the invidious labeling of compensatory education pupils and their segregation from classes of 'regular' pupils". (p.47). There is not space here to detail the arguments of Glass and Smith, nor to review the evidence they have mustered against "pull-out" procedures in Title I. Suffice it now to note that the arguments against pull-out procedures likely will influence the future of Title I programs. With regard to early childhood Title I programs, however, the potential impact of the arguments against pull-outs seems likely to be limited largely to post-kindergarten levels, for only at those levels do pull-out procedures seem prevalent.

## 2.7 Summing Up:

What are the implications of these trends for early childhood Title I programs and for efforts to evaluate these programs? It is of course impossible to be certain, but in summing up this chapter let us try to suggest something of what these trends may hold for the future of early childhood programs funded under Title I. First, the continuing high level of interest in early childhood education and the increasing enrollments of preschool children in educational programs suggests that use of Title I funds to support early childhood programs may be on the increase. The continuing concern for parental involvement indicates that evaluations of early childhood programs should take this phenomenon into consideration, although different forms of parent involvement likely would require different forms of evaluation. Increasing coordination for ECE at the state level suggests that patterns of organizing and funding ECE programs may be changing. Hence it may be increasingly difficuly clearly to isolate Title I ECE programs from other publicly supported ECE programs. The continuing emphases on comprehensive services and the more recent concern for ecological approaches to ECE suggest that early childhood programs funded under Title I will continue to have multiplicity of goals--unlike many Title I programs for older children--and that there will be increasing efforts to coordinate such programs with other public services to disadvantaged families. Recent worry over basic skills achievement suggests, however, that concern for comprehensive services may be tempered somewhat in the future, with priority given to services that seem to afford most promise of boosting basic skills achievement.

The increasing emphasis on individualization--which is one of the strongest trends apparent in our review--implies that the very idea of a program as a set

of procedures or activities administered alike to some set of pupils may be on the wane. Evidence of the trend toward individualization is widespread and it has special relevance for Title I programs. In the Education Amendments of 1974, Congress encouraged "where feasible, the development for each educationally deprived child participating in a program under (Title I) of an individualized written educational plan (maintained and periodically evaluated), agreed upon jointly by the local education agency, a parent or guardian of the child, and when appropriate, the child." (P.L. 93-380, Sec. 131). The trend toward mainstreaming and the arguments against Title I pull-out programs also suggest that it may be increasingly difficult in the future to isolate discrete program treatments which traditionally have been the subject of educational program evaluation.

These suggestions offer little clear direction to solving the problems of evaluating Title I ECE programs. In the main, they point to the difficulty of doing so. However, rather than elaborating more on these issues here, let us turn in Chapter III to consider broader issues concerning evaluations of other federal early childhood programs. In Chapter IV, we will focus attention more closely on Title I and see how the broad trends identified here show up on Title I.

## III. EVALUATION OF FEDERAL PROGRAMS IN EARLY CHILDHOOD EDUCATION

The upsurge of interest in early childhood education in the 1960's was accompanied by a fundamentally new development--an explicit national commitment of federal aid to education in general and to the education of young children in particular. As noted in Chapter II, federal investment in preschool programs had increased to the amount of $1 billion annually in the early 1970's. A 1970 report entitled Federal Programs for Young Children (Child Development Staff, 1970) identified more than 200 federally funded programs serving young children. Paralleling the rise in federal funding of such programs has been an increasing concern for evaluating the effectiveness of these programs. Since 1970, the federal government has invested tens of millions of dollars in evaluating programs for young children, and millions more have been invested by state and local governments in evaluating selected federal education programs (GAO, 1977).

The purpose of this chapter is to review the recent literature on federal evaluations of early education programs in order to identify basic problems encountered in conducting these evaluations. The review focuses on identification of methodological and procedural problems encountered rather than the substantive findings of these evaluation efforts. We anticipate that these problems are ones that must be confronted in devising a system for evaluating Title I ECE programs.

Before launching into the text of our review, four limitations of the review should be pointed out. First, the review is based largely on secondary sources - that is, on documents encompassing several individual studies. The constraints on the time and resources available for this review preclude a full consideration of primary source materials - that is, individual evaluations of particular early childhood programs. Second, the chapter

focuses on past problems in conducting evaluations of federal early childhood education programs, not on successes. This focus is adopted for several reasons. The problems are simply more prominent than successes. Also, successes, when they have been won, often seem to turn up in ways that were unanticipated at the start of evaluations--as we shall see in section 3.5. We pay more attention to problems of past evaluations of federal ECE programs than to successes simply because they are more relevant to a future system for evaluating ECT I programs then are successes of such large-scale evaluation endeavors. Nevertheless, the reader should keep in mind that the focus on problems leads to a rather bleak and perhaps somewhat incomplete picture of past federal evaluation efforts. Third, this chapter does not treat evaluations of Title I programs. These will be dealt with separately and in more detail in Chapter IV. Fourth, the chapter does not discuss tests and instruments for young children. This topic arises with such frequency in early childhood evaluation and is of such prominent concern in our feasibility study that a separate chapter (V) will be devoted to it.

## 3.1 Broad Goals and Narrow Evaluation

Perhaps the most pervasive problem apparent in federal evaluations of early childhood programs is the discrepancy between the broad goals of such programs and the narrow scope of program evaluations. This problem has been a familiar one at least since 1969, when the Westinghouse-Ohio evaluation of Project Head Start was published. In their critique of the Westinghouse evaluation, Smith and Bissell (1970) noted that most Head Start centers had no one particular goal, but instead had multiple goals, "aimed at children's physical, mental and social well-being in general." Hence, Smith and Bissell found it "difficult to believe that 'overall' policy decisions about increasing, decreasing or removing funds from Head Start can be justified on the basis of a study that focused only on children's performance on a few cognitive and affective tests."

In response to Smith and Bissel's critique, Cicirelli, Evans and Schiller (1970) defended the Westinghouse evaluation as imperfect, but still a useful contribution to policy-making.

> In the real world of government social action programs, decisions are going to be made either in the presence or absence of information. If the social scientist hopes to increase at all the rationality of these decisions he will often have to depart from ideal evaluation requirements. Only rarely if ever, for example, will it be possible in the evaluation of ongoing social action programs to arrange for the random assignment of subjects to experimental and control groups; and time pressures will often make it necessary to forego longitudinal measurements. Of course, such compromises involve risks and the greater the compromises the more likely it is that results will be in error. But in our opinion these risks outweigh those of the only other alternative we have, which is to continue to make decisions mainly on the basis of faith, testimonials, and power politics (pp. 125-6).

Nevertheless, the discrepancy between broad goals and narrow evaluations has continued to trouble Head Start research and evaluation efforts. In their review of Head Start research since 1969, Mann et al (1977) reported:

> The most serious gap that was observed was the failure of most of the research to evaluate Head Start as a comprehensive service program. Often a narrow definition of objectives was used which failed to look at the "total" program. In particular, program goals concerning child health and community change were overlooked.

A similar pattern was found in a review of national evaluation studies of project Follow Through (Haney, 1977; Villaume and Haney, 1977). Follow Through (FT) was designed as a comprehensive program with a variety of goals. Millions were spent in the early years of FT on evaluating a wide range of those goals, but over the years the national FT evaluation narrowed to focus in the end on a minority of the goals originally set out for Follow Through. Although the official rules and regulations for FT suggested seven possible criteria by which the program might be evaluated, the national evaluation focused mainly on one; namely, comparison of FT students' cognitive and affective

development as measured on standardized tests with that of non-Follow Through comparison students. In this respect the Follow Through evaluation was broader than many program evaluations; children's affective development as well as their cognitive and academic skill development were covered in the FT test battery. Nevertheless, goals of the program such as parent involvement, comprehensive services and institutional change received scant attention in this major evaluation effort. Though there were many reasons why the FT evaluation effort focused on children's cognitive and affective development, as measured by standardized tests, (e.g., insturmentation problems, political and bureaucratic pressures and cost considerations) two emerge as most important. First was an implicit assumption behind evaluation planning that program goals such as parent involvement, service provision and institutional change were important only insofar as they contributed to children's educational development. Hence, went the apparent rationale, if these program elements were successful, they ought to yield increased test scores. The second major reason for the focus on children's test scores was that evaluation of this sort fit well with the prevailing paradigm of evaluation--namely, the treatment-control group model of evaluation research. Facets of the FT program like parent involvement and institutional change simply did not fit very well with this model of evaluation research.

The contrast between broad goals and narrowly focused evaluation extends far beyond Head Start and Follow Through. A 1973 review of all federal programs for young children noted that "evaluation measures were primarily limited to the cognitive realm, to IQ and achievement tests" (White et al, 1973, Volume IV, p. 20). Similarly, a review by Hertz (1977) of the impact of federal early childhood programs on children observed:

> [M]uch of the evidence on the impact of Federal programs pertains to children's cognitive development and performance. It is apparent, however, that even conclusions about effects on cognitive development must be regarded as tentative and incomplete, especially in light of the narrow range of measures used (p.19).

In sum, it is apparent that this is an old and prevalent problem. Evaluations of early childhood programs invariably seem to be far narrower than the broad goals generally held out for such programs. The problem is clearly one which developers of a system for evaluating Title I early childhood programs must confront. In part, the problem stems from the complexity of most early childhood interventions--a complexity in which the rationale for various program elements is either unclear, or so extended that it is hard to reduce to the form of simple propositions amenable to empirical investigation. The problem derives also from the heavy reliance of evaluation research on the traditional treatment-control model of experimental research, often relying solely on standardized test instruments as outcome measures. Hence, it seems likely that in order to treat broader aspects of early childhood interventions, a system for evaluating Title I early childhood programs cannot rely solely on the treatment-control model of evaluation.

### 3.2 Long-range Goals versus Short-range Evaluation

The ethos of social reform, of which compensatory education is one manifestation, generally holds that interventions should aim to help people not just with their proximate problems, but more ambitiously with their long-range prospects. So it is with federal programs for disadvantaged children. The Follow Through program, for example, ambitiously aims at "aiding in the continued development of children to their full potential" (U.S. Congress, 1967).

There is widespread discrepancy between the long-range rhetoric of such program goals and short-term reality of most program evaluation. White et al (1973) in their review of federal programs for young children pointed out the short-range horizon of most program evaluations:

> Most evaluations have measured the effects of projects over one year only. At times, this has led to judging projects successful when, over the course of several years, they would not be so judged. Or, it may not accurately indicate the possibility of projects that involve major organizational changes (which might depress achievement initially). (p. 20).

Mann et al (1977) in their review of Head Start research since 1969 discovered a similarly myopic view in most assessments of Head Start effects. They found that only seven out of fifty-nine studies reviewed looked at long-term effects of Head Start even though, as the reviewers pointed out, the issue of long-term effects would be of increasing importance as Head Start enters its second decade. In his wide-ranging review of federal early childhood programs Hertz found a similar pattern:

> Research on Home-Start, Parent Child Centers, Sesame Street, the Electric Company, the Handicapped Children's Early Childhood Education Program and day care, adds little to our understanding of the long-term impact of early intervention since appropriate follow-up testing has not been undertaken (Hertz, 1977).

In the national evalaution of Follow Through a longer perspective was adopted. Follow Through was evaluated not just over a single school year, but instead over the full three- or four-year span of the program. Even so, the evaluation has been criticized as too short, ignoring what happens to children after they leave the program. Yet, despite much discussion of fade-out effects (where apparent program effects disappear after children leave the program) and sleeper effects (long-term effects not evident until after children leave the program), the national Follow Through evaluation has focused mainly on end-of-third-grade test scores of children who had been in the program for three or four years, and has not employed follow-up testing (Haney, 1977).

The duration of evaluation is closely related to another issue troublesome to evaluators--the attrition of program participants out of intervention programs. The problem is generally not so severe in evaluations spanning

only a single year, but in long-range evaluations the problem can become quite serious. For example, in Project Follow Through, it was found that approximately half of the initial program participants dropped out of the program every two years. Hence, at the end of the four-year program, only around one-quarter of the original program participants were still around to be tested. So the attrition problem was a real hindrance to evaluators trying to generalize results to the whole of the population initially served and to the full four-year program.

It is clear that the attrition problem can be overcome in follow-up evaluations if the investigator is prepared to make special efforts. For example, in their unusual study of the persistence of preschool effects, Lazar, et al. (1977) were able to track down two-thirds of their original sample eight to fifteen years after sample subjects left the preschool programs evaluated. Their remarkable follow-up effort was, however, rather expensive. It cost Lazar et al. around $250 to track down and test each of their follow-up subjects. Such costs may not be exorbitant for a research project like that of Lazar and the Developmental Continuity Consortium, but for a regular annual evaluation of a program like Title I, such follow-up costs clearly would be prohibitive.

In addition to the practical problems of longitudinal evaluation, there are difficult theoretical problems also. After children leave an intervention program, their experiences are likely to be quite diverse. Moreover, long-term educational experiences of children who participate in a special early education program may differ in important respects from the long-term experiences of children with whom their educational progress might be compared. In the jargon of educational experimentation, long-term program effects are likely to be confounded with post-treatment experience.

These sorts of considerations led Rand researchers conducting a feasibility study for a national evaluation of Head Start to recommend against a longitudinal evaluation. They wrote:

> Our recommendation is that if a national evaluation is conducted, it should not be a longitudinal evaluation. The Office of Child Development was originally interested in this possibility. However, we advise against it for several reasons: serious theoretical problems with linking outcomes produced by Head Start and outcomes at a later stage in the child's development; cost of following up a nationally representative sample for three or four years; diminished cell sizes when the data are controlled at least for the child's school, and possibly for his classroon. We also argue that Head Start is appropriately held accountable for only immediate, not long-term, effects.
>
> We assume that the development of a child is a somewhat Markovian process--i.e., outcomes at $time_n$ are primarily determined by immediately preceding outcomes at $time_{n-1}$, not so much by those at $time_{n-2}$ or $time_{n-3}$. For example, if a child has reading problems in third grade, it is probably more appropriate to hold the second grade responsible than Head Start. (Raizen and Bobrow, 1974, pp. vii-viii).

In sum, the long-term consequences of early childhood Title I programs are of real interest, but practical and theoretical considerations raise doubts as to the feasibility of conducting longitudinal impact evaluations on a regular basis. While specially funded longitudinal evaluations (of the sort conducted by Lazar, et al.) might be carried out, the prospects for making rigorous longitudinal evaluations a regular part of an ongoing Title I evaluation system are not promising.

### 3.3 Program Definition and Implementation

Another problem apparent in virtually all reviews of federal early education programs is the widespread failure of evaluations to describe adequately the intervention being evaluated and to assess whether or not the intervention ostensibly being evaluated actually was implemented as intended. White et al. (1973) for example, noted that "Project descriptions,

fundamental to our project categorizations and to our ability to relate project characteristics to outcomes, were often vague and general." (p. 20). In his 1977 review of impact evaluations of federal early childhood programs, Hertz noted the same problem and recommended:

> More attention must be given to implementation in future studies, and not just as a confounded variable, but as the focus of research. Just as we need to know which models are more effective, we need to know which ones can be easily and successfully implemented, why they can be, and how they can be. (p. 10)

The widespread failure of program evaluations to describe programs and assess their implementation was noted too in the Social Research Group (SRG) review of Head Start research since 1969. The SRG investigators found that only a small minority of Head Start evaluations clearly specified program inputs. In consequence, they noted, many questions regarding differential implementation were left unanswered (Mann, et al., 1977).

In the national evaluation of Follow Through the implementation issue was a major focus of investigation. One of five major questions addressed in the FT evaluation was "Do differences in implementation of an educational strategy bear any relationship to the differences in impact? (Stebbins, et al., 1977, p. 7). Despite a variety of efforts to answer this question, it remains as one of the major unanswered questions of this multimillion dollar evaluation effort. For an account of the various efforts to assess the implementation of Follow Through models, see Haney, 1977, pp. 151-169. Although the Follow Through evaluation has long been billed as a planned variation experiment, no one has yet come up with a way to assess the implementation of FT's planned variations. In fact, even the descriptions of the Follow Through models remain more like statements of philosophy than clearcut descriptions of

classroom practice, curricula or even educational objectives for each grade-level encompassed in the Follow Through program. See, for example, Stebbins, et al (1977). More complete descriptions of FT models are available in forthcoming USOE volume of FT sponsors' own evalautions. However, even these descriptions remain remarkably vague.

The failure of so many federal evaluations of early childhood programs to describe the interventions they presume to evaluate seems highly anomalous. The rationale for most educational program evaluations is that they will serve to identify techniques useful in helping to improve student learning. In the idiom of experimental research, evaluation aims at identifying program treatments that lead to desired outcomes. Much has been written in recent years about the manner in which evaluators can efficiently estimate treatment effects, but relatively little attention seems to have been given to what constitutes an adequate description of the "treatments" being evaluated. This situation is anomalous since estimations of effects, however precise or valid they may be, hold little if any value if one cannot identify the interventions that produced them.

How to describe adequately the nature of educational interventions and the extent of their implementation is a problem that troubles far more than merely early childhood program evaluations. Nevertheless, it is a problem that any system for evaluating early childhood programs must face for it is hard to envision what purposes such an evaluation system might serve if causes as well as effects are not adequately described. As we will see in Chapter IV, what little we know of early childhood Title I programs suggests that they encompass a variety of very different kinds of interventions. Their variety may simply preclude describing them all in terms of any fixed set of dimensions.

### 3.4 Uncertain Findings

Another issue apparent in most reviews of federal evaluations of early childhood programs is that not much at all is known for certain as a result of these efforts. For example, White et al , (1973) in their massive review of federal programs for young children noted that all conclusions concerning early elementary education interventions were conditioned by several broad qualifications. Similar qualifications were noted with respect to preschool interventions, and with respect to day care interventions the reviewers were even less sanguine:

> The most reasonable conclusion about existing data for early day care would seem to be that the data are limited, preliminary and inconclusive. (White, et al., 1973, p. 25).

Four years later in his review of the impact of federal early childhood programs, Thomas Hertz reached a similar conclusion:

> Answers to questions about program impact must be pieced together from fragmentary findings, with the result that many questions cannot be answered at all, and most answers that can be provided must be considered as tentative and incomplete (Hertz, 1977, p.1) .

With respect to Follow Through--which has easily been the most thoroughly researched and evaluated federal intervention project for young children--the following was observed:

> [T]here is no easy or mechanical way to sum up the findings from the various national evaluation reports [on Project Follow Through]. Different analyses have been based on different samples, have encompassed different spans of the program, and have used widely varied analytical techniques. It is impossible even to sum up the summaries of the major evaluation reports...[Therefore, the process of summarizing findings] is like fitting available pieces of a puzzle together to show the outlines of a picture. Instead of yielding definitive findings, it can give us only probable conclusions (Haney, 1977, p. 307).

The latest national evaluation study of Project Follow Through yielded a total of ten specific findings (Abt, 1977, pp. 135-155). Yet nine of

those specific findings have already been challenged (House, et al., 1977, 1978; Haney, 1977). What is the single "finding" of this multi-million dollar evaluation that has gone unchallenged?

> The effectiveness of each Follow Through model varied substantially from site group to site group; overall model averages varied little in comparison (Abt, 1977, p. 135).

The prevalence of uncertain findings in the research and evaluation literature on federal programs for young children is sobering. The studies encompassed in the reviews cited above were not just regular program evaluations. Rather, they generally were specially funded efforts, often conducted over lengthy periods of time and staffed by professionals with specialized training in educational research and evaluation. In light of the paucity of clear-cut findings from such endeavors, it seems highly doubtful as to the level of certainty to which an ongoing system for evaluating early childhood Title I programs can reasonably aspire.

### 3.5 Rare Use of Evaluation Findings

The point usually noted as distinguishing program evaluation from basic research is that it is keyed to the decision-making process. Valid information regarding program effectiveness, it is presumed, can inform decision-makers regarding future programs. Nevertheless, the link between program evaluation and decision-making has not been often or easily forged. As the latest annual evaluation of Office of Education programs noted: "While the design and conduct of good evaluations is difficult, costly, and time consuming, the application of evaluation findings is an even harder thing to accomplish" (OBPE, 1977, p. 18).

In light of the prevalent problems in evaluations of federal early education programs noted above it is not surprising that the connection

between evaluation findings and program decision-making has been a tenuous one. Consider for example, the case of the national Follow Through evaluation. Millions of dollars were spent to evaluate Follow Through and literally dozens of formal evaluation reports on Follow Through have been completed over the last ten years. Yet in the summer of 1977, when the Joint Dissemination Review Panel (JDRP) of the Office of Education and the National Institute of Education met to judge the evidence of effectiveness of Follow Through projects, data from the Follow Through evaluation was considered in a way that was contrary to the whole rationale behind the Follow Through planned variation experiment. The purpose of the Follow Through planned variations was to develop and evaluate alternative models of early childhood education. In accordance with this rationale, the bulk of effort in the national Follow Through evaluation was directed to estimating the effectiveness of alternative Follow Through models. Yet for a variety of reasons the JDRP review of Follow Through focused not on FT models, but instead on individual FT projects. So ironically, even though the whole purpose of the FT evaluation had been to find effective models, the evaluation data did not contribute to decision-making about models. For more information on the relationsip between the Abt evaluation and the JDRP review of Follow Through see Haney (1978).

The direct use of evaluation research in decision-making regarding Head Start also seems to have been quite rare. For example, although the Westinghouse-Ohio evaluation of Head Start in 1969 received widespread publicity, that evaluation study seems to have had little direct impact on decision-making. Several articles in the journal Evaluation in 1973, for example, concluded that the Westinghouse evaluation had little impact. In a more recent assessment Datta (1976) concluded that there were several kinds of

consequences flowing from the Westinghouse findings. Datta suggested that the Westinghouse evaluation did in fact: (1) promote specific change in Head Start programming (away from summer programs, and more toward full-year programs); (2) spur experimentation with alternative models of Head Start programs (specifically with Head Start planned variations); (3) lead to considerable debate and reconsideration of early childhood education policy; and (4) promote widespread re-examination of evaluation techniques and policies. In short although "the direct effects [of the Westinghouse evaluation] on practictioners are probably negligible," Datta maintains that the study did in fact have a variety of uses in changing the general climate of opinion and the manner in which a variety of observers thought about early childhood education and evaluations of educational interventions (Datta, 1977, pp 158, 175-6).

This phenomenon is by no means unique to early childhood program evaluation. In a wideranging review of the impact of social science research on educational policy Cohen and Garet (1975) noted that influence of research on policy seems far less direct and clearcut than is often envisioned in most models of information-based decision-making. They concluded:

> 1. Most policy-oriented research, at least in education, tends to influence the broad assumptions and beliefs underlying policies, not particular decisions.
> 2. Better methodology and policy relevance in applied research in education have not produced more convergent findings. This is in part because most policy-oriented research concerns programs with broad and conflicting aims, but it is also attributable to methodological conflict among research approaches and to the fact that the advance of applied research tends to complicate and redefine issues. As a result, improving applied research does not tend to produce more authoritative advice about social policy.
> 3. The justification for applied social research carried out for the state cannot rest on the idea of instrumental rationality, because applied research, at least in education, has not significantly increased the objective information base for decisions. (pp. 39-40)

The reasons for the lack of direct use of evaluation findings in program decision-making are varied. In part the cause may lie in the fact that evaluation research rarely can yield comprehensive information on all issues relevant to policy-makers. As Mann et al., (1977) noted in their review of Head Start research "the research often omitted dimensions of crucial importance to policy formulation. Only one study [of the 59 reviewed] looked at the cost/benefit of the program." (p. 20).

More broadly, however, it is vital to recognize that decision-making never takes place solely on the basis of empirical evidence concerning program processes and effects. Inevitably it is conditioned by political bureaucratic, historical and value considerations. On this point, Datta states:

> The weakest link in evaluation...is not in strengthening design, measures, data collection and analysis although dispute may focus on these, and strengthening is often needed. It is rather in anticipating the range of decisions within the power of the many decision makers who may apply what is learned from evaluations (Datta, 1977, p. 177).

Datta's challenge is clearly relevant to any system designed to evaluate early childhood Title I programs. Who is it that might be informed by such a system? In what sorts of decisions might they make use of information yielded? What sorts of evaluation information might inform those decisions?

These questions have no easy answers. They are important with respect to program evaluation in general, but seem especially important with regard to Title I evaluation. As we will see in Chapter VI, the Title I evaluation literature reveals some large discrepancies between information needs apparent at the national, state, and local levels of the American educational system, as well as discrepancies in information needs identified by different sources at the national level.

3.6 Summing Up

What do these problems imply for a system to evaluate early childhood Title I programs? The discrepancy between the generally broad goals of early education programs and the narrow focus of most evaluations of them suggests that the typically broad goals of early childhood interventions may not be easily encompassed in any single evaluation reporting format. In theory, the solution to program diversity might be simply to acquire more information on all programs, so that data collection encompasses information pertinent to the full range of programs. From a practical point of view, however, such a solution appears problematical. As a previous Title I evaluation study noted, "the meaningfulness, accuracy and validity of the reported data would be inversely related to the amount requested" (Gamel, et al, 1975, p. 7). As a classic paper on the topic of institutional information gathering put it, "there are often good organizational reasons for bad clinical records," especially when much information is required (Garfinkel, 1967). In other words, a workable solution to the broad goals and diversity of early childhood programs would not seem to be simply to require more information from all programs. Instead, a diversified system allowing differential attention to different goals for different programs may be more appropriate.

Both practical and theoretical problems in longitudinal evaluation suggests that despite the long-term goals of most early childhood interventions, evaluations of them may need to be more modest and avoid long-term impact studies, at least on any regular yearly basis. Problems of program definition and implementation suggest that a system for evaluating early childhood Title I programs might give more attention to descriptive questions of what these programs really are. The uncertainty of findings from past large-scale evaluations of federal early education programs reinforces the suggestion that the goals of an

ECT I evaluation system ought to be relatively modest in terms of experimental rigor and may simply have to settle for uncertainty with respect to many questions of itnerest. The rare use of past findings from early childhood evaluations suggests also that an ECT I evaluation system should be aimed at clearly specified information needs. In chapter VI we return to the issue of information needs that an ECT I evaluation system might serve. However, before turning to that issue, in the next chapter we will review the history of past Title I evaluations and in chapter V review the state of the art of early childhood testing and instrumentation.

## IV. EARLY CHILDHOOD TITLE I PROGRAMS

What sorts of early childhood programs have been funded under Title I and what do we know about them? This is the question addressed in this chapter of our literature review. In trying to answer the question, we draw on four sources of information. First are past national evaluations of Title I programs. Second are Fiscal Year 1976 State Evaluation Reports on Title I. Third are reports on exemplary educational programs. Fourth is information on early childhood education Title I programs uncovered in a search of the ERIC system.

The chapter is organized into seven sections. The first summarizes what we have learned about early childhood Title I programs from our review of past national evaluations of Title I programs in general. The second section pulls together information on the scope of early childhood Title I programs as compared with the overall Title I effort. The third recounts what we know about the kinds of early childhood programs funded under Title I. The fourth discusses the issue of how children are selected to participate in early childhood Title I programs. The fifth section discusses the special role of parents in such programs. The sixth section describes the special problems apparent in state evaluations of Title I programs in general and in evaluations of early childhood programs in particular. The seventh and final section sums up the chapter and suggests some implications of our findings for our future work.

Before we launch into the text of this chapter, a reminder on terminology may be useful. Recall that in this chapter the term "early childhood programs" is used to refer to programs serving children in prekindergarten, kindergarten and first grade age levels. Specifically excluded from this review are programs serving only children above the first grade. Nevertheless, as will become evident, it is not always possible to differentiate programs serving

PK-1st grade children from those serving older children.

## 4.1 Past National Evaluations of Title I Programs

Over the past decade, almost a score of national studies of Title I have been carried out. A 1977 literature search revealed that the ERIC system contains more than 1700 smaller studies concerning Title I. Nevertheless, this vast literature has not wrought a very large store of wisdom. Michael Timpane (1976) in summing up U.S. Office of Education efforts to evaluate Title I programs from 1965 to 1972 termed them a "wave of failures". As observed in the latest effort to synthesize Title I evaluations, "the first five years of evaluation of Title I were essentially a total loss in terms of achieving any of the valid objectives for evaluation". (McLaughlin, Gilmartin & Rossi, 1977, p.10). So, a detailed review of the Title I evaluation literature would not be a promising endeavor. Moreover, such an undertaking would not be possible within the time and resource constraints of the current project. Hence, all we attempt to do in this section of Chapter IV is to survey eighteen major efforts to evaluate Title I programs with an eye to what attention they have given to early childhood Title I programs. These eighteen investigations are listed in Table 4.1. The first fifteen evaluation studies have been summarized in Rossi et al (1977). It should be noted that many of these investigations encompassed not only Title I programs, but more broadly compensatory education efforts in general.

Two observations regarding early childhood programs can be drawn from the past large-scale efforts to evaluate compensatory education programs; one obvious and the other not so obvious. First, and obvious, is that relatively

Table 4.1: Attention Given to Early Childhood Programs in National Evaluations of Title I*

| STUDIES | YEAR(S) | ATTENTION TO PK, K, AND GRADE 1 PROGRAMS |
|---|---|---|
| Studies Summarized in Rossi et. al., (1977) | | |
| 1. TEMPO Study (Mosbach) | 1967-68 | None to PK or K. Data from only one of 11 districts included in final report Analysis included grade 1 |
| 2. 1967-68 Survey (see Rossi et al) | 1968-70 | None. Focuses on grades 2,4, & 6. |
| 3. 1968-69 Survey (Glass) | 1969-70 | None. Focuses on grades 2,4, & 6 |
| 4. It Works (Hawkridge) | 1968-71 | High proportion of successful projects identified included PK, K or grade 1 |
| 5. CPIR Surveys (NCES) | 1970-73 | No special attention, but includes some data from PK, K and grade 1 levels |
| 6. Washington Research Project | 1969 | No direct attention |
| 7. Exploratory Study (McLaughlin) | 1970-71 | Study concentrated on full-year programs for grades 1,2, and 3. |
| 8. Literature Review (Gordon) | 1970-71 | Gives some attention to infant and preschool programs |
| 9. 5-Year Synthesis (Wargo) | 1971-72 | Little direct attention but includes some data and descriptions of PK, K and grade 1 programs. |
| 10. Compensatory Reading Study (Trismen, OBPE) | 1971-76 | Focuses on reading programs for students in grades 2,4, and 6 |
| 11. Planar Reports (Planar) | 1972-73 | Indirect use of PK, K and grade 1 data from FY 72 SEA Reports |
| 12. State Report Review (Gamel et al) | 1974-75 | Indirect use of PK, K and grade 1 data from FY 1969-74 SEA Reports |
| 13. State Report Review (Thomas & Pelavin) | 1974-76 | Data analysis limited to samples from grades 1-12 from SEA Reports. No attention to PK or K |
| 14. GAO Report (1975) | 1974-75 | No direct attention. |
| 15. PIPS Evaluation (Stearns) | 1974-77 | Study of implementation of first six PIPs, of which one is for K-6 students, one for first graders |
| More Recent Studies | | |
| 16. Longitudinal Analysis (Pelavin & David) | 1976-77 | Lowest grade level invluded in data analysis in grade 3 |
| 17. AIR Reviews (McLaughlin et al, & Rossi) | 1976-77 | No direct attention, but does include info. from other sources regarding PK, K and grade 1 programs. |
| 18. NIE Studies (Hill, Frechtling, Syracuse) | 1976- | National survey has info on grades K-8. Effects study includes grades 1 & 3 |

* Full citations for reports from each of these major studies are provided in the bibliography.

little attention has been given to programs in the early childhood range--that is, programs below the second grade level. The first national evaluation of Title I in 1967 and 1968, commonly termed the TEMPO Study, gave only scant attention to early childhood programs. The national surveys of Title I programs in 1967-68 and 1968-69 focused solely on grades 2, 4 and 6, as did the Compensatory Reading Study, which lasted from 1971 to 1976. The CPIR surveys (NCES, 1976) in the early 1970's gave early childhood programs no special attention, nor did Wargo's five-year synthesis (1972) or McLaughlin, et al's exploratory study (1971). Also, reviews of Title I programs by the Washington Research Project/NAACP Legal Defense Education and Education Fund in 1969 and by the Government Accounting Office in 1975 gave no particular attention to Title I early childhood programs. Recent reviews of state Title I evaluation reports by the Planar Corporation (1973), Gamel, et al (1975), and Thomas and Pelavin (1976) gave brief attention to such programs since some state reports did include data broken down by Pre-Kindergarten, Kindergarten and grade 1 levels. The Thomas report in 1976 included first grade data, but excluded prekindergarten and kindergarten data.

More recent studies, too, have given only passing attention to early childhood programs. The special analyses of longitudinal data from compensatory education programs by Pelavin and David (1977) included data from only the third grade upward. The A.I.R. studies in 1977 gave early childhood programs no direct attention, but did include a little information from other sources concerning prekindergarten, kindergarten and first grade programs. The NIE Survey of Compensatory Education estimated in 1976 that nationwide, 38% of Title I school districts provide preschool/kindergarten

readiness activities funded wholly or in part by Title I, but beyond that estimate provided no more direct information on such programs (Hill and Rotberg, 1976) Another part of the NIE Compensatory Education Survey provided data on the effects of Title I first grade services, but gave no information whatever on PK and K services (Frechtling, 1977). So in large measure, we simply do not know what kinds of services have been provided by Title I below the second grade.

The second and less obvious conclusion to be drawn from this literature review concerns the type of study in which early childhood programs do receive relatively more attention. In only two of the 18 studies listed in Table 4.1 do early childhood programs figure very prominently: in the Hawkridge (1968) "It Works" reports and in Stearn's (1977) evaluation of Program Information Packages or PIPS. Also, preschool and early elementary programs show up prominently in portions of the A.I.R. reviews that summarize information on exemplary programs (Rossi, et al, 1977, p. 62; McLaughlin et al, 1977, pp. 42-44). In short, early childhood programs figure prominently in each of the reports dealing with exemplary programs. In fact, most of the exemplary programs treated in these studies were originally identified in the Hawkridge, "It Works" studies. Table 4.2 lists, for example, the 19 programs which Rossi et al in 1977 found most convincing among the 42 "It Works" programs originally identified by Hawkridge et al in 1968 and 1969. Of the nineteen programs listed, eight are aimed primarily at preschool age children and three others at least include first-grade students. In other words, roughly half of the exemplary programs which Rossi et al found most convincing serve children in the prekindergarten to first grade range.

A similar though not quite so striking prominence to early childhood programs emerges from USOE's 1976 publication, Education Programs That Work.

Table 4.2: Rossi, et al.'s "Most Convincing of the Successful Programs Identified in the [It Works] Study"

| Grade Level | Title | Specific Ages Or Grades | Site |
|---|---|---|---|
| Preschool | Learning to Learn | 5 Yrs. | Jacksonville, FL. |
| | Preschool Program | 3-5 Yrs. | Fresno, CA |
| | Program Early Push | 4 Yrs. | Buffalo, NY |
| | Preschool Program | 3-4 Yrs. | Oakland, CA |
| | Early Childhood Project | PK-Gr. 3 | New York, NY |
| | Perry Preschool Project | 3-4 Yrs. | Ypsilanti, MI |
| | Diagnostically Based Curriculum | 5 Yrs. | Bloomington, IN |
| | Academic Preschool | 4-5 Yrs. | Champaign, IL |
| Elementary | Programmed Tutorial Reading | Gr. 1 | Indianapolis, IN |
| | The Plus Program | | Buffalo, NY |
| | Afternoon Remedial and Enrichment Program | Grds.3-8 | Buffalo, NY |
| | Intensive Reading Instructional Teams | Grds.3-6 | Hartford, CN |
| | After School Study Centers | Grds.2-6 | New York, NY |
| | More Effective Schools | Elem. Grds. | New York, NY |
| | The School and Home Program | Elem. Grds. | Flint, MI |
| Elementary/ Secondary | Homework Helper Program | Grds.3-6 | New York, NY |
| Secondary | Junior High Summer Institutes | Grds.7-9 | New York, NY |
| | Project R-3 | Grds.8-9 | San Jose, CA |
| | College Bound Program | Grds.9-10 | New York, NY |

Source: Adapted from Rossi, R.J., McLaughlin, D.H., Campbell, E.A., and Everette, B.E. Summaries of Major Title I Evaluations. (Palo Alto, CA: July, 1977) p. 62. "Grade level" categorizations are those given by Rossi et al. Specific grade levels are those noted in McLaughlin et al (1977, pp. 42-44).

Of 151 active exemplary education programs listed in that book, 39 are classified as early childhood/parent readiness programs, and many others serve the whole kindergarten-to-grade-six or grades K-6 elementary age range. Again, it appears that early childhood programs show up with unusual frequency in identifications of exemplary programs.

An obvious question arises. Why? There are no clear-cut answers, but several seem plausible. First, it may be that numerous exemplary early childhood programs have turned up simply because special attention has been given in the last decade or so to developing high quality programs for children in this age range. This seems plausible, because as noted in chapter II, interest in early childhood programs seems to have risen steadily over the last ten to fifteen years. A second and related possibility is that a large number of early childhood programs may have been considered by investigators searching for exemplary programs. Therefore more exemplary early childhood programs may have been discovered simply because more were considered. We have no direct information on this possibility. A third possibility is that early programs more often than later programs are identified as having demonstrable impact because of the plasticity of early development, the notion discussed in Chapter II. Again, we have little direct evidence on this point. For now we note only that the simple notion of early plasticity versus later fixity, though popular in the 1960's, has recently been challenged by several developmental psychologists (Hunt, 1975; Kagan, 1977). Later, in section 4.6, we will see how this notion turns up in recent Title I literature.

Another possible explanation for the number of "exemplary" early childhood programs has to do with the nature of the intervention. It may be that early childhood intervention and particularly preschool programs represent sharper departures from the status quo (or no treatment condition) than is the case for

later interventions. After all, after age five, almost all children receive regular doses of some sort of education. If this is so, early interventions may be identified as exemplary more often than later grades programs simply because larger interventions likely yield demonstrable impact more often than smaller interventions. In section 4.3, we will offer some evidence on the relative magnitude of early versus later compensatory interventions--or at least Title I program interventions.

A fifth possible explanation for the frequent appearance of early childhood interventions among exemplary programs has to do with the means by which program effects are assessed. It may be that the measurement techniques by which exemplary programs were identified were somehow biased in favor of early interventions. In chapter V we will elaborate on this possibility. For now, let us simply sum up what we have learned from our review of past national Title I evaluations.

Two points are clear from this portion of our literature review. First, almost all of the national evaluations of Title I over the past decade have given no more than passing attention to Title I programs in the preschool to grade-one range. Second, the only place in the national evaluation literature where early childhood programs show up with any regular prominence is in efforts to identify exemplary programs--that is, ones with demonstrable effectiveness. There seem to be several possible explanations for the unusual prominence of early childhood programs among the exemplars. We have little direct evidence on which of these explanations is most likely. However, in following sections of this paper we will discuss information bearing on some of them.

## 4.2 The Scope of Early Childhood Title I Programs

In the last section, we learned that early childhood compensatory education programs do exist and that they seem effective in unusual frequency. Beyond this, however, the large-scale efforts to evaluate Title I provided little information on the scope or nature of such programs. In this section, we pull together what information we have gathered from various sources on the scope of Title I early childhood programs.

In attempting to do this, we again run into the problem that early childhood, as defined for this project, does not correspond to the manner in which that term is used elsewhere. Also, most sources simply do not disaggregate information on Title I in ways compatible with the given scope of our investigation. For example, recall that the NIE Survey of Compensatory Education estimated that 38% of the nation's school districts provide services defined as "preschool/kindergarten readiness activities", funded wholly or in part by Title I (Hill and Rotberg, 1976, p. III-21). If grade 1 activities were to be added in with prekindergarten to kindergarten activities, this figure would presumably be much higher, but data disaggregated in this manner is simply not available from the NIE survey nor from any other nationwide Title I study.

Another source of information on the scope of early childhood Title I programs is performance and accounting reports annually submitted by SEA's to USOE. Table 4.3 summarizes information from FY 1975 and FY 1976 state performance and accounting reports. Once again, however, it should be noted that our efforts are somewhat stymied by the manner in which the data available to us are reported. These data have been summarized by USOE according to the following grade-level spans: prekindergarten-to-kindergarten, grades 1-6,

and grades 7-12 levels. Hence, we must settle for numbers and proportions of Title I participants in PK-K levels as our best estimate of the scope of Title I early childhood programs.

Table 4.3 shows that nationwide, approximately eight per cent of Title I participants are in PK-K programs. This amounts to around 400,000 students participating in Title I PK-K programs. Emphasis on PK-K Title I programs varies widely across the states, however In several, less than one per cent of Title I students were in PK-K programs. The state apparently placing most emphasis on early childhood Title I was Wisconsin--with 27 per cent of its Title I children in PK-K programs. In terms of absolute numbers, the state with the most children in Title I PK-K programs doubtlessly was California, although the FY 1976 performance and accounting report for California is not available.

These conclusions are roughly corroborated by data in Table 4.4, which summarizes the numbers of participants in state Title I programs at the PK, K and first grade levels. These data have been gleaned from our review of FY 1976 state Title I reports. We must warn very strongly, however, against placing overmuch faith in these data since there is literally no single type of information which could be unambiguously summarized across annual state Title I evaluation reports. For several states, FY 1976 reports are simply not available. In cases where reports are available, report formats, manner of data aggregation, and definition of terms used all vary so widely that virtually no data are uniformly reported in all state reports.

TABLE 4.3: TITLE I PARTICIPATION BY STATE FOR FISCAL YEARS 1975 AND 1976

PRE-KINDERGARTEN - KINDERGARTEN AND TOTAL (PRE-K TO GRADE 12)

| STATE | FISCAL YEAR 1975 | | | FISCAL YEAR 1976 | | |
|---|---|---|---|---|---|---|
| | TOTAL (PRE-K - 12) | PRE-K - K | %PRE-K - K | TOTAL (PRE-K - 12) | PRE-K - K | %PRE-K - K |
| TOTALS | 4,524,530 | 360,602 | .08 | 5,055,353 | 404,483 | .08 |
| Alabama | 146,866 | 7,960 | .05 | 143,170 | 9,187 | .06 |
| Alaska | 5,191 | 860 | .17 | 5,501 | 1,015 | .18 |
| Arizona | 41,299 | 2,762 | .07 | 56,097 | 4,561 | .08 |
| Arkansas | 68,429 | 1,282 | .02 | 70,223 | 1,848 | .03 |
| California | 601,968 | 85,245 | .14 | NA | NA | NA |
| Colorado | 36,646 | 2,090 | .06 | NA | NA | NA |
| Connecticut | 36,650 | 447 | .01 | 51,193 | 8,108 | .16 |
| Delaware | | | | 9,364 | 1,038 | .11 |
| Florida | 122,619 | 7,504 | .06 | 148,787 | 11,369 | .08 |
| Georgia | 141,430 | 13,809 | .10 | 122,721 | 16,117 | .13 |
| Hawaii | 8,237 | 627 | .08 | 10,234 | 791 | .08 |
| Idaho | 18,579 | 1,037 | .06 | 14,862 | 124 | .01 |
| Illinois | 171,880 | 16,566 | .10 | 169,902 | 16,375 | .10 |
| Indiana | 107,909 | 8,159 | .08 | NA | NA | NA |
| Iowa | 60,966 | - | - | 56,313 | 1,665 | .03 |
| Kansas | 34,335 | 2,758 | .08 | 33,407 | 1,398 | .04 |
| Kentucky | 114,792 | 5,326 | .05 | 112,757 | 4,971 | .04 |
| Louisiana | 127,317 | 7,351 | .06 | 150,402 | 10,309 | .07 |
| Maine | 30,058 | 2,629 | .09 | 31,373 | 2,113 | .07 |
| Maryland | NA | NA | NA | 62,055 | 12,445 | .20 |
| Massachusetts | 77,129 | 10,706 | .14 | 67,057 | 9,278 | .14 |
| Michigan | 127,172 | 19,593 | .15 | NA | NA | NA |
| Minnesota | 58,207 | 6,411 | .11 | 71,354 | 4,400 | .06 |
| Mississippi | 139,942 | 1,081 | .01 | 121,429 | 1,171 | .01 |
| Missouri | 76,840 | 4,736 | .06 | 83,312 | 5,363 | .06 |
| Montana | 7,893 | 60 | .01 | 11,313 | 85 | .01 |
| Nebraska | 35,737 | 1,809 | .05 | 31,445 | 953 | .03 |
| Nevada | 2,446 | - | .00 | 5,502 | 120 | .02 |
| New Hampshire | 6,415 | 225 | .04 | 7,000 | 150 | .02 |
| New Jersey | 92,464 | 11,889 | .13 | 103,508 | 15,005 | .14 |
| New Mexico | 22,868 | 4,614 | .20 | 28,441 | 4,422 | .16 |
| New York | 503,350 | 25,689 | .05 | 419,540 | 10,005 | .02 |
| North Carolina | 133,161 | 11,554 | .09 | 129,899 | 9,558 | .07 |
| North Dakota | 14,856 | -- | .00 | 14,081 | 75 | .01 |
| Ohio | 130,361 | 13,069 | .10 | 132,938 | 13,331 | .10 |
| Oklahoma | 79,496 | 140 | .00 | 94,155 | 180 | .00 |
| Oregon | 40,867 | 2,284 | .06 | 47,400 | 2,962 | .06 |
| Pennsylvania | 251,208 | 13,284 | .05 | 279,371 | 18,755 | .07 |
| Rhode Island | 10,872 | 71 | .00 | 25,252 | 288 | .01 |
| South Carolina | 144,876 | 5,110 | .04 | 133,124 | 3,512 | .03 |
| South Dakota | 19,600 | 2,177 | .11 | 18,814 | 898 | .05 |
| Tennessee | NA | NA | NA | 122,550 | 6,963 | .06 |
| Texas | 317,633 | 32,739 | .10 | 408,765 | 39,853 | .10 |
| Utah | 15,752 | 1,531 | .10 | 18,283 | 2,332 | .13 |
| Vermont | 12,660 | 1,366 | .11 | 13,055 | 1,306 | .10 |
| Virginia | 109,161 | 2,928 | .03 | 100,987 | 4,742 | .05 |
| Washington | 79,505 | 3,316 | .04 | 68,685 | 3,531 | .05 |
| West Virginia | 42,915 | 40 | .00 | 42,913 | 38 | .00 |
| Wisconsin | 53,137 | 14,749 | .28 | 61,973 | 16,469 | .27 |
| Wyoming | 5,577 | 224 | .04 | 3,523 | -- | .00 |
| District of Columbia | 17,000 | 2,550 | .15 | 18,400 | 3,220 | .18 |
| American Samoa | NA | NA | NA | NA | NA | NA |
| Guam | 1,245 | -- | .00 | 1,250 | -- | .00 |
| Puerto Rico | NA | NA | NA | 234,812 | 6,574 | .03 |
| Trust Territory | 17,269 | NA | .00 | 7,987 | 20 | .00 |
| Virgin Islands | 1,735 | 245 | .14 | NA | NA | NA |

Source: "Fiscal Year 1975 Performance Reports - Total Number of Children Participating" and "ESEA Title I - Number of Children Participating by Grade Level - Fiscal Year 1976", Summary Tables of the Annual State Performance and Accounting Reports, prepared by Division of Education for the Disadvantaged, copies obtained from Jim Ogura. Excludes participation state institutions. NA means data not available.

TABLE 4.4: SUMMARY OF REVIEW OF FY 1976 STATE TITLE I EVALUATION REPORTS: NUMBERS OF PARTICIPANTS IN PK, K AND G-1 PROGRAMS

| State [1] | Pages [2] | No. Participants [3] Pk | K | 1 |
|---|---|---|---|---|
| Alabama | 58 | - | 9 | 22 |
| Alaska | 120 | - | - | - |
| Arizona | 27 | 1 | 4 | 5 |
| Arkansas | 56 | 1 | 1 | 5 |
| California | 48 | - | 161 | 170 |
| Colorado | 111 | .8 | 1 | 3 |
| Connecticut | 85 | 4 | 4 | 7 |
| Delaware | 200 | ND | ND | NB |
| Florida | 65 | → | 11 | ND |
| Georgia | 80 | → | 10 | 10 |
| Hawaii | 51 | .2 | .5 | 1 |
| Idaho | 49 | - | .1 | 2 |
| Illinois | 345 | ND | ND | ND |
| Indiana | 200 | | 5.5 | ND |
| Iowa | 51 | - | - | - |
| Kansas | 55 | .5 | 1 | 4 |
| Kentucky | 69 | - | 5 | 15 |
| Louisiana | 42 | 2 | 9 | 14 |
| Maine | 13 | - | 2 | 4 |
| Maryland | 38 | 2 | 11 | 14 |
| Massachusetts | 120 | 4 | 7 | 13 |
| Michigan | 29 | 8 | ND | ND |
| Minnesota | 26 | ? | ? | ND |
| Mississippi | 31 | → | 1 | 16 |
| Missouri | 55 | 1 | 4 | 5 |
| Montana | 19 | - | .1 | 1 |
| Nebraska | 48 | → | 1 | 11 |
| Nevada | 194 | - | - | - |
| New Hampshire | 36 | - | ? | - |
| New Jersey | 25 | ND | ND | ND |
| New Mexico | 25 | - | 4 | ND |
| New York | 119,30 | - | - | - |
| North Carolina | 85 | - | 14 | ND |
| North Dakota | 36 | - | ND | ND |
| Ohio | 24 | 6 | 7 | 22 |
| Oklahoma | 35 | ND | ND | ND |
| Oregon | 50 | .5 | 3 | 6 |
| Pennsylvania | 72 | - | 19 | ND |
| Rhode Island | 61 | .3 | 2 | 3 |
| South Carolina | 36 | → | 4 | ND |
| South Dakota | 21 | - | 1 | 2 |
| Tennessee | 63 | → | 7 | ND |
| Texas | 32 | 5 | 39 | 52 |
| Utah | 39 | - | 2 | 3 |
| Vermont | 50 | - | ? | - |
| Virginia | 88 | 1 | 4 | 23 |
| Washington | 45 | → | 4 | ND |
| West Virginia | 25 | - | - | ND |
| Wisconsin | 131 | 7 | 10 | 9 |
| Wyoming | 28 | - | - | .3 |
| District of Columbia | 108 | - | 4 | 4 |
| American Samoa | | | | |
| Guam | | | | |
| Puerto Rico | 85 | - | 4 | ND |
| Trust Territory | | | | |
| Bureau of Indian Affairs | 42 | - | - | - |
| Virgin Islands | 27 | - | - | - |

Footnotes

1. There are a total of 57 "states" from which reports are potentially available. From three such units (American Samoa, Guam and Trust Territory), no reports were available. From five states (Georgia, Massachusetts, Michigan, New Jersey, and West Virginia), only FY 1975 reports were available. Complete citations for the reports from which this information was gleaned are given in Appendix I.

2. Several state reports are unpaginated. In such cases, page length has simply been estimated. In many cases, reports contain main texts plus separately numbered appendices. Listed here are the total length of the reports, with both texts and appendices included.

3. Estimated here are the reported number of participants (in thousands) at prekindergarten, kindergarten and first grade levels. All numbers greater than 1,000 have been rounded to the nearest thousand. Numbers below 1,000 have been rounded to the nearest hundred and expressed as tenths of 1,000. A dash (-) indicates that there is no clear indication in the relevant report of children participating in a particular grade level. ND means not disaggregated and refers to cases in which there is indication of children participating at particular grade levels, but numbers are not disaggregated by grade levels. Arrows in the PK column indicate cases in which PK-K have not been disaggregated in state reports. Question marks refer to cases in which numbers of participants are not clearly indicated but where there is strong reason to presume that there are students participating at the indicated grade levels.

Nevertheless, drawing on these two sources of uncertain information, we can come up with a list of twenty-one states in which there seems to be substantial activity in early childhood Title I programs:*

| | |
|---|---|
| Alabama | Michigan |
| Alaska | New Jersey |
| California | New York |
| Connecticut | Ohio |
| Georgia | South Dakota |
| Illinois | Texas |
| Kentucky | Utah |
| Louisiana | Vermont |
| Maryland | Virginia |
| Massachusetts | Wisconsin |

District of Columbia

---

* A state is included in this list if more than 10% of its Title I participants in either FY 1975 or FY 1976 were in PK-K programs as summarized in Table 4.3 or, if the state's FY 1976 annual evaluation report indicated that at least 20,000 children participated in Title I programs in the PK-grade 1 range.

## 4.3 Nature of Early Childhood Title I Programs

What sorts of prekindergarten, kindergarten and first grade programs have been funded under Title I? This is the question addressed in this section. However, before trying to answer this question, let us comment briefly on two recurring problems: sources of information and terminology.

The prime sources of information available for describing early childhood programs are FY 1976 state Title I evaluation reports and descriptions of exemplary Title I programs. In the remainder of this chapter, we will draw extensively on the state Title I evaluation reports. Instead of salting the text of the chapter with references to these reports, we have simply listed full references to all of the FY 1976 state reports in the bibliography appended in this report (see appendix I). Hence, in the text of the paper we provide only page references for direct quotations from the state reports.*

In describing early childhood programs we again face the problem of terminology. The terms early childhood, preschool, kindergarten, and child development are frequently used in state Title I evaluation reports. Unfortunately, however, they are used in widely different fashion in different reports. For example, in Connecticut "early childhood" is used to refer to kindergarten-to-third-grade programs, not including pre-kindergarten programs. In Colorado, in contrast, "early childhood" seems to be used to refer only to prekindergarten programs. In some state reports,

---

*Note that as recounted in the footnote #1 in Table 4.4, only FY 1975 reports were available for five states (Georgia, Massachusetts, Michigan, New Jersey and West Virginia). Also, for three state units (American Samoa, Guam and Trust Territory), no reports at all were available.

preschool is used synonymously with kindergarten, while in others kindergarten is identified separately, both as a grade level and as an instructional level. In the pages that follow we will attempt to use the terms prekindergarten, kindergarten and first grade as we defined these terms in chapter I. When these terms are used in any other manner we will footnote this usage.

Prekindergarten

Most state reports provide little description of programs funded or services provided at this level. There seem, however, to be two basic types of programs funded at the prekindergarten level--home-based and center-based programs. The home-based program most clearly identified in state evaluation reports is the Toy Resource program in Connecticut.

> Eleven L.E.A.'s [in Connecticut] conducted Toy Resource programs during the 1975-76 school year, serving a total of 428 pupils. In Connecticut Toy Library Resource programs, children between 2 1/2 years and 4 years of age are visited weekly in their homes over a 10 month period by a trained toy demonstrator. In the presence of one or both parents, use of a toy to teach a basic cognitive concept (size, shape, color, number, e.g.) is demonstrated. The toy is left for the parent to use with the child until the demonstrator's next visit. (p. 26).

Though the issue is not entirely clear, center-based prekindergarten programs appear to be more prevalent than home-based ones. Unfortunately, none of the state reports provides very clear descriptions of center-based prekindergarten programs. Nevertheless, one such program in Cedar Rapids, Iowa has recently been validated by the Joint Dissemination Review Panel (JDRP) as an exemplary project. The Cedar Rapids Title I program is called "Pre-kindergarten Education for the Disadvantaged Child" and aims at alleviating

> developmental lags and deficiencies of selected prekindergarten children in target school areas by providing professional screening and diagnostic procedures to identify each child's needs, and then designing an individualized program to remediate these needs. Parental

> involvement in all phases of the Early Childhood program is encouraged at home as well as in the center, thus assuring understanding and follow-through of prescriptive activities....Classes meet four days per week, Monday through Thursday, for two and one-half hours per day. Each child participates in approximately 600 minutes of classroom time per week. Friday of each week is devoted to program planning and review, staffings with resource personnel, parental involvement, and inservice training.*

Other early childhood programs serving prekindergarten children seem to combine aspects of home- and center-based programs, for example, in an exemplary Title I program in Chicago, Illinois, entitled "Child-Parent Centers Activity" (USOE, 1976, p. 44).

Available descriptions of prekindergarten Title I programs indicate that the goals sought by such interventions are generally broader than reading and math achievement, goals which usually are the focus of elementary school Title I programs. For example, New Jersey reported needs assessment of Title I children by grade level, summarized across all LEA's in the state. At the pre-kindergarten level seven needs were noted.

- General School Readiness Skills
- Language Concept Development
- Socialization
- Motor Coordination
- Math/Reading Readiness
- Oral Communication
- Visual Motor Coordination (p. 11)

Kindergarten

At the kindergarten level, Title I early childhood programs seem to focus somewhat more narrowly on what are commonly called school readiness skills. For example, Kentucky reported that the most commonly sought

---

*"Exemplary Project Proposal, Cedar Rapids Community Schools, Cedar Rapids, Iowa." This program is not included in USOE's Education Programs That Work, apparently because it was approved by the JDRP after this volume went to press in 1976.

objectives of Title I kindergarten programs in that state were "general readiness skills as measured by published readiness tests, locally constructed readiness tests, checklists, etc." (p. 11).

State Title I evaluation reports provide no more than passing descriptions of the sorts of kindergarten programs funded under Title I. Descriptions of exemplary Title I projects again provide us with an idea of what these programs are. For example, the Baptist Hill kindergarten program in Greenville, Alabama was approved by the JDRP in 1974.

> The goal of this project is to provide experience necessary for school readiness to five-year-old children of low-socioeconomic background. Children experience success daily, interact with peers and adults, and learn to respect the rights of others. Experiences are provided to help children increase their vocabulary, learn to speak in sentences, communicate more effectively, grow in independence and initiative, understand and deal with the world around them, and develop problem-solving skills. Learning centers are set up in each classroom. Student needs are assessed and continuous evaluations made to ensure program effectiveness. The program is monitored by a parent involvement component. (USOE, 1976, p. 43.)

Another exemplary Title I project in Cincinnati, Ohio provides full-day kindergarten to disadvantaged children:

> The program aims to strengthen the child's perceptual and auditory skills, improve his/her language development, and develop the potential of the family to motivate the child's learning. Teachers complete skills observation checklists for each child. This data along with Boehm test results are used to group students for work with the Sullivan Reading Readiness materials on language and concept development. Six to eight headphone units and one central console in each classroom facilitate individual or small-group work. Perceptual skills are sharpened through use of manipulative materials. Inservice for teachers and aides is continuous, approximately one-half day per month. (USOE, 1976, p. 48.)

In some states, emphasis on Title I kindergarten programs is clearly increasing. For example, in Alabama, the number of Title I kindergarten children is reported to have doubled between 1974 and 1976. In Wisconsin, the percentage of districts offering Title I kindergarten programs has steadily

increased over the past decade; from near zero in 1968, to 40 per cent in 1972 and to more than 50 per cent in 1976 (p. 49).

In several states, however, Title I funding of kindergarten programs has decreased in recent years. For example, New Mexico reports that a phase-in of statewide kindergartens is replacing Title I kindergarten programs. "It is anticipated that by spring of 1978 all Federally funded kindergarten programs [in New Mexico] will be absorbed by state funds." (p. 17). North Carolina reports that there is a diminution of Title I kindergarten programs "due to the new inclusive State program in kindergarten." Also, Virginia reports that kindergarten activities previously provided under Title I "are now provided division-wide from other funds" in 43 Virginia school divisions (p. 35). In these instances, it appears that Title I has acted as a catalyst to stimulate state or local funding of kindergarten programs.

First Grade

In the Title I evaluation literature, first grade programs are distinguished separately from later grade-level programs far less frequently than prekindergarten or kindergarten programs. Nevertheless, there is considerable evidence that more children are involved in first grade Title I programs than in either prekindergarten or kindergarten programs. As indicated in Table 4.4, among those states for which FY 1976 data were disaggregated by grade-level, all but one had more Title I participants in first grade than in kindergarten. As shown in Table 4.5, as best we can tell from past SEA reports approximately twelve per cent of all Title I participants are first-graders, as compared with only around eight per cent in kindergarten programs.

Table 4.5: Percentage of Title I Participants in Prekindergarten, Kindergarten and First Grade Levels Presented by Year, as Reported in SEA Reports, 1971-1974.

| | FY 71 | FY 72 | FY 73 | FY 74 |
|---|---|---|---|---|
| Prekindergarten | 1.8 | 2.5 | 1.8 | 1.1 |
| Kindergarten | 8.5 | 8.3 | 7.6 | 7.8 |
| Grade 1 | 12.3 | 13.1 | 13.5 | 12.5 |
| Number of States Reporting Data Disaggregated by Grade-level | 13 | 14 | 18 | 16 |

Source: Adapted from N. N. Gamel, et al. State ESEA Title I Reports: Review and Analysis of Past Reports and Development of A Model Reporting System and Format (Mountain View, CA: October 1975), p. 65.

The manner in which first-grade data are usually lumped together with data from other grades in Title I reports represents more than arbitrary data aggregation. Virtually all Title I programs serving first-graders encompass more than just that one grade level. Such programs generally serve a range of students; typically kindergarten-to-third-grade, first-to-third-grade or even first-to-sixth-grade.

This pattern of programs serving multiple grade levels is evident in the early childhood Title I projects that have been passed by the JDRP. Altogether 45 Title I projects have earned exemplary status by being approved by the JDRP. Twenty-nine of the 45 projects serve children in the prekindergarten-to-first-grade ranges, at least in part. Eleven of the exemplary Title I projects serve kindergarten children; two of these serve kindergarten exclusively. Six of the exemplary Title I projects serve prekindergarten children; two do so exclusively. In contrast, of the 21 exemplary Title I projects serving first grade, only one does so exlusively. It is the Programmed Tutorial Reading project developed in Farmington, Utah.

> The objective of the Programmed Tutorial Reading project is to provide tutoring to underachieving 1st graders in beginning reading as a supplement to conventional classroom teaching. The teaching strategy employs many elements of programmed instruction: frequent and immediate feedback, specified format, and individualized pace. However, whereas programmed instruction has often sought errorless or near-errorless learning with many cues at first, followed by a fading of cues, the tutorial program proceeds in the opposite manner with minimal cueing at first, followed by increased prompting until the child can eventually make the correct responses. Through the use of 11 different Item Programs, the tutoring behavior of each tutor is carefully controlled. Project components are systematically programmed so that any decision made

> by a tutor is limited to judging the correctness of a reading response or the appropriateness of an answer to a question. Reinforcement is an important part of the instructional strategy. (USOE, 1976, p. 119)

In short, Title I programs serving first-graders rarely seem to do so exclusively. Far more often than not, such programs also serve later grades students.

Also, like Title I services to later grades, services to first-graders generally seem to be quite narrowly focused on boosting children's achievement in language and mathematics. For example, New Jersey reported that LEA's most frequently identified the following special needs at the first-grade level.

| | |
|---|---|
| Language Experience: | Work attack/knowledge |
| | Comprehension |
| | Vocabulary |
| | Phonetic Analysis |
| Computational Skills: | Concepts |
| | Computations |
| | Problems/Application (p. 10) |

The New Jersey report identified precisely these same special needs as most common also for grades 2, 3, 4, 5, and six.

* * * * * * * * *

In summary, the nature of Title I services at the three early childhood levels, at least insofar as we can tell from available descriptions, is substantially different. At the prekindergarten level, Title I funds are more often used to fund entire intervention programs, often with both home- and center-based components. To some extent Title I also funds whole programs at the kindergarten level, although in several states other public money is being used to fund kindergarten programs begun with Title I funding.

The whole-program nature of Title I prekindergarten and kindergarten services is reflected in cost data presented in several FY 1976 reports. Data in reports from Alabama, Georgia, Louisiana, North Carolina and Wisconsin indicate that per pupil costs of programs serving these levels are substantially higher than those for later grades Title I services. In North Carolina, for example, per pupil expenditures for kindergarten were $712, whereas corresponding costs for reading and mathematics services were respectively $298 and $205 per pupil (p. 11). The higher costs for PK and K Title I programs appear to stem mainly from the fact that at these levels Title I funds frequently pay for entire programs instead of merely for supplements to already existing school programs. This interpretation is supported too by cost data indicating that, in comparison to later grades program costs, food and transportation budget items represent an unusually large proportion of preschool and kindergarten costs in Alabama (p. 11), Georgia (pp. 28, 35), and Tennessee (pp. 31, 36).

The differences in the goals and nature of Title I programs at the prekindergarten, kindergarten, and first grade levels have clear implications for evaluations. We will elaborate on some of these in a later section of this chapter. For now, simply recall the suggestion made earlier. It may well be that prekindergarten programs show up frequently in lists of empirically judged exemplary projects simply because of the nature of the intervention. Compensatory services at the prekindergarten level seem to amount to much sharper interventions into the status quo than later grades services, which usually are added on in supplementary fashion to pre-existing school services.

## 4.4 Selection and Needs Assessment

Title I is a program to help educationally disadvantaged children. The official rules and regulations governing the administration of Title I require that children be selected to receive Title I services only after a multi-stage process, involving designation of Title I attendance areas, ranking of attendance areas on the basis of educational need, identification of all target-group children, identification of children most in need, and finally, selection of children identified as having the greatest need (DHEW Federal Register, September 28, 1976).

These provisions seem to have resulted in widespread use of standardized tests to identify children's educational needs. South Carolina, for example, reported that standardized tests are used alone or in part as the criteria for selecting Title I children in more than 80 per cent of that state's districts (p. 15). Arizona reported that standardized test results were used in 97 per cent of that state's districts as a method of assessing children's educational needs (p. 13). (Teacher recommendations were a source from which to determine educational need in 96 per cent of the Arizona districts.) Table 4.6 shows the methods which LEA's in Iowa most frequently reported using to identify students' educational needs, with test scores and teacher recommendations again ranking most prominently as needs assessment devices.

Selecting children for early childhood Title I programs raises particular problems. Group-administered standardized tests (the most commonly reported means of selecting Title I participants) are simply not appropriate for use with very young children, as will be discussed in

TABLE 4.6

Frequently Used Criteria for
Identification of Children in Need of
Supplementary Educational Programs and Services*

| Rank | Method of Identification | Frequency of Response |
|---|---|---|
| 1 | Achievement test scores | 401 |
| 2 | Teacher recommendations | 393 |
| 3 | Record of unsatisfactory academic achievement | 354 |
| 4 | Readiness scores | 287 |
| 5 | Evidence of psychological and/or social maladjustment | 244 |
| 6 | Indication of inadequate family resources | 198 |
| 7 | Referrals from outside agencies | 109 |
| 8 | School Drop-Outs | 28 |

* Most agencies use a combination of methods as indicated by frequency scores.

Source: Iowa State Title I Evaluation Report, FY, 1976, p.3.

chapter V. So education agencies are forced to rely on other procedures. Indeed, the issue has proved so troublesome that it recently caused an administrative disagreement between USOE's Division of the Disadvantaged (DED) and the HEW Audit Agency. In the past DED had taken what officials term a "soft" position on selection of preschool children into Title I programs.* They maintained that since it is usually unfeasible to test young children and since preschool children almost by definition had no past school records from which educational need might be judged, that LEA's could select children for Title I preschool programs simply on the basis of residence within Title I attendance areas. In 1975, however, in the course of reviewing one LEA's administration of Title I funds, the HEW Audit Agency ruled that the district was in violation of federal regulations because it was not selecting preschool participants on the basis of educational need. The ruling was disputed by DED but the conflict was recently resolved in favor of the Audit Agency's interpretation by the USOE General Counsel. According to the ruling, participants for Title I prekindergarten and kindergarten programs, like participants for later grades programs must be selected on the basis of educational need. Hence, it is likely that selection of Title I preschool participants will become an increasingly important issue in the administration of such programs.

Unfortunately, few state Title I reports disaggregate information on selection procedures at different grade levels. One of the most extensive discussions of preschool selection procedures was provided by Alabama:

> [Preschool] participants were tentatively identified by Headstart records, commercial kindergarten records, economic status, [and] records of siblings who had been unsuccessful in school. Participation in the program

---

*This account of controversy over selection of preschool children is based on a communication with Mr. James Ogura, USOE Division of the Disadvantaged, November, 1977. The term preschool is used here to mean both prekindergarten and kindergarten.

> required pre-testing to determine eligibility. Those children who ranked in the lowest stanines (and percentiles) were selected first -- those in higher stanines were selected on a space available basis. (pp. 38-39)

In contrast to this elaborate procedure, Alaska reports that in several Alaskan villages, "All three, four and five-year-old children are eligible to attend Early Childhood Development Programs in their villages"(n.p.) In what appears to be a more intermediate position, South Carolina reports that Title I kindergarten programs in that state use "free or reduced price lunch eligibility only as a selection criterion."(p.15).

At the first grade level, use of test results as selection criteria was apparently the most common method (e.g. in District of Columbia and Utah), as is clearly the case generally for Title I programs. One agency, however, expressed strong reservations about the adequacy of tests scores for this purpose.

> Perhaps the most important finding obtained from teachers' responses...was that a substantial number of eligible students are not being identified by the norm-referenced achievement test. Conversely, a number of students not needing Title I services are, on the basis of faulty test scores, being placed in the Title I program. The exclusive use of norm-referenced tests for determining Title I eligibility is generating considerable undue dollar and human costs. Although no selection process will be one hundred percent accurate, the present system, adopted under USOE recommendation, is inadequate and should be supplemented by teacher judgment. (District of Columbia, p. 5)

In short, the issue of how to select children with the greatest educational needs to receive Title I services is a terribly important issue. Also, it bears strongly on the issue of how to evaluate Title I services. As noted in USOE's User's Guide to the Title I evaluation models, the manner of student selection is a major determinant of the appropriateness of using particular models for evaluating Title I programs in the grade 2 to grade 12 range (Tallmadge and Wood, 1976).

## 4.5 Parental Involvement

Legal requirements for parent participation in Title I have increased sharply over the past decade. Regulations now call for Title I Parent Advisory Councils (PACs) to be established for each Title I school. Some states now also require LEAs to report on parental involvement in PACs, at least in terms of numbers of PAC meetings held and attendance at those meetings.

Evidence of increasing attention to parent involvement in Title I early childhood programs is apparent in documents entered in the ERIC system. In November, 1977, we carried out a search of the ERIC system to identify all documents described as pertaining to both Title I and to preschool.* The search yielded a total of 40 documents from the period 1966 to 1969; 36 from 1970 to 1973,; and 23 in the 1974-1977 period. In the first period only two, or 5% of the documents were described as pertaining to both parent and preschool education. In the second period, seven, or 20%, of the documents were identified as pertinent to parent education, and in the 1974-1977 period, six or 26% of the Title I preschool documents also dealt with parent education.

Despite this apparent increase in concern for parent involvement, the degree of commitment to parental involvement in different states, at least as expressed in FY 1976 state reports, seems to vary widely. On the low end of the enthusiasm scale was the Virgin Islands, which reported that the Virgin Islands PAC "as in previous years, did not get off the ground" (p.4). Other state reports evidenced somewhat more enthusiasm for parental involvement. Yet, the degree of importance attributed to it was not always strong. For example, when Virginia LEAs were asked to rank factors contributing to Title I success, "parent support"

---

*The ERIC search was designed to retrieve all documents which had as descriptors "ESEA Title I" and one of the following: "preschool children," "preschool education," "preschool evaluation," or "preschool programs." Of the 99 documents, most were local school district Title I evaluation reports (62) or state Title I evaluation reports (30). Why the number of documents entered into ERIC pertaining to Title I preschool programs has decreased since 1966 is unclear.

came out sixth, after:

1) use of teacher aides;

2) use of more and better educational equipment and supplies;

3) reduced size of classes and teacher load;

4) in-service training; and,

5) use of specialists and consultants (p.25)

Perhaps the strongest expression of support for parent involvement in Title I came from Illinois:

> Parents must fully understand the purpose of Title I, and their role in helping to fulfill that purpose. During FY 1976, the majority of LEA's were in compliance with the federal regulations on Parent Advisory Councils. However, it is imperative that efforts to involve parents in Title I be continued. It has been proven that student achievement is affected by parental interest, and the success of Title I is at least partially dependent on the effectiveness of the PAC's. (p.8)

Kentucky reported data showing that parent involvement is far more extensive in that state's kindergarten components (pp. 4-6) than in later grade programs. Parent involvement is also obviously of essential importance in prekindergarten home-visit programs like Connecticut's Toy Resource Project. Nevertheless, Connecticut pointed out that parental involvement seems to have been slighted in terms of evaluation. "Only six of the fifteen programs which cited parental involvement as a major program goal reported an evaluation of that goal" (p. 27). The issue of how much parents are actually involved in Title I programs was also raised recently in the NIE compensatory education study. A synthesis of the Syracuse Research Corporation (1977) study of Title I administration reported that "attention to, and understanding of, administrative responsibilities in the area of parent involvement was minimal in most of the school districts visited."

In sum, the federal mandate for parental involvement in Title I PAC's is clear. Yet the degree of support for parent involvement in different states across the country seems widely varied. This contrast, coupled with the obvious importance of parental involvement in early childhood Title I programs (especially home-based programs) clearly suggests that a system for evaluating early childhood Title I should take this program component into account. Likely, there will be a need to differentiate different forms of parent involvement in such an evaluation system, but parental involvement does seem to have special relevance for early childhood Title I programs.

## 4.6 Evaluation of Early Childhood Title I Programs

In Section 4.1 we noted that Title I early childhood programs seem to have appeared in unusual frequency in nationwide efforts to identify exemplary compensatory education efforts. In this section, we will turn from such large scale evaluation efforts to see how these programs have been evaluated in FY 1976 state Title I Evaluation Reports. In the first few paragraphs, we offer some general comments on the annual state Title I evaluation reports. Then we focus more narrowly in the remainder of this section to describe how the state Title I reports treat early childhood programs.

Even a cursory look at the FY 1976 state Title I evaluation reports reveals that there is little information that can be unambiguously aggregated in straightforward fashion across more than a handful of state reports. Our experience in reviewing the 1976 reports has been much the same as that of Gamel et al who reviewed state Title I evaluation reports from 1971 to 1974:

> It was clear long before the formal analysis was completed, that the typical State Title I Annual Evaluation Report falls far short of the ideal. A substantial number of the reports which were examined contained no data which could be aggregated, and the sum total of the reports could not be used to synthesize a national-level assessment of Title I effectiveness. (Gamel et al, 1975, p.50)

The most common method of reporting Title I outcome data was in the form of grade equivalent gains; either in terms of average monthly gains (AMG) or grade equivalent gains during entire school year. Several states noted that although the grade equivalent metric has its weaknesses, it is still desirable for use because it is "well understood." Another state reported that it was continuing to use the grade equivalent metric both because it is well understood and because it is "valid".

Most state reports do not indicate the degree to which reported outcome data are representative of all children participating in Title I. Nevertheless, there is considerable information to suggest that reported test score data is

not representative of all Title I children even within particular states. For example, Idaho reports that test data is available for less than half of the participants in that state. (p.25). From South Dakota, the state report indicates that test data was not reported when gains were "excessive." (p. 14). And in New York, data were excluded for just the opposite reason - because they showed that "participant groups made no gains while receiving Title I services." (p.78).

The degree of influence exerted by states on local Title I reporting seems to vary widely. Minnesota, in what appears to be one of the stricter forms of state direction, instructed LEAs as follows:

> For reporting reading scores, use only one of the nationally standardized reading tests that was included in the Anchor Test Study. There were:
>
> California Achievement Test CAT
> Comprehensive Test of Basic Skills CTBS
> Gates MacGinitie Reading
> Iowa Test of Basic Skills ITBS
> Metropolitan Achievement Test MAT
> Science Research Associate Achievement Test SRA
> Sequential Test of Educational Progress STEP
> Stanford Achievement Test SAT
>
> For reporting mathematics achievement, the districts are permitted to use any nationally normed standardized test. For grades pre-Kg. through 2, the district may still use teacher judgment. (p.15)

In Illinois, in contrast, the SEA apparently recommended no particular outcome measures. The report from that state consists entirely of data concerning the degree to which LEAs were able to meet objectives established at the local level. Utah, in another example of a more laissez-faire approach to local reporting, noted that "the key indicators (i.e. tests and teacher rating) that were used to determine student gain and growth were selected by district personnel and local parent advisory councils" (p.7).

Several states evidenced skepticism as to whether aggregation of Title I

evaluation results was very meaningful. For example, Maine reports:

> It should be noted that the Maine Title I, ESEA, unit does not consider the data being collected nationally for Title I, ESEA, evaluation purposes as having a high statistical reliability in assessing success of Title I, ESEA program participants. The bulk of current testing, whether norm or criterion referenced, is not designed to measure performance of these students. The national samples exclude and the norm tables extrapolate data relative to this highly vulnerable and quixotic group.
>
> The best measure of success for this group of students is personal observation by teachers and trained adults who work directly with them on as close to a one-to-one basis as possible. It is impossible, therefore, to report the most significant data on gain and performance. (p.8)

And, in a section of its report on product evaluation, Iowa comments:

> It remains difficult, however, to present [product evaluation] data into meaningful state totals and/or averages. The better programs and the poorer programs, according to achievement data presented, become lost in the statewide totals and/or averages. In essence, the evaluation reports tend to be quite meaningful at the local school district level, less meaningful at the state level, and quite possibly even less meaningful at the national level. (p. 11)

Two state reports also evidenced concern over the costs of evaluation procedures. Illinois reported:

> [Many] LEAs use standardized tests to measure the effectiveness of Title I on their students academic performance. These tests are costly to buy, administer and score. As a result, many LEAs use tests that are available and/or less expensive, but not always appropriate. (p.8)

Without elaboration, but one suspects not without purpose, Kansas reported the following:

> In an attempt to identify the costs associated with Title I evaluation, participating LEA's were asked to provide the following cost estimates for evaluation activities:

| | |
|---|---|
| Cost of Testing Materials | $113,216 |
| Cost of Staff Time (including teacher time) to administer, score, record pre-tests and post-tests. | $424,428 |
| Cost of staff time to prepare evaluation report (to SEA) | $ 86,724 |
| Cost for services of local project evaluations (exclusive of costs listed above) | $ 67,705 |
| TOTAL | $692,703 |
| Amount of above total funded with Title I monies | $527,670 = 67% |

The above Title I expenditure estimate for evaluation represents 4.3% of the total fiscal year Title I allocation of $12,269,304. (p. 37)

Despite such concerns, or perhaps in part because of them, several states seem to be moving to implement one or more of the RMC models for evaluating Title I programs and to report results in terms of the NCE metric. In their FY 1976 reports, eighteen states evidenced knowledge of, or intention to implement one of the RMC models or to report results in terms of the NCE metric.* Nevertheless, like all else, attitudes and enthusiasm for the new reporting procedures seem to vary widely across the states. South Dakota, for example reports:

> After carefully considering the NCE system, the problems and deficiencies with grade equivalence and the degree of commitment to the change that existed at the federal level, the SEA decided to take "initial steps to install the system." (p.7)

The New Hampshire SEA commented:

> This office has mixed feelings about this effort, but we are attempting to utilize some of the techniques devised by RMC...

* Alabama, Arizona, Arkansas, California, Connecticut, Florida, Maine, Montana New Hampshire, New York, Pennsylvania, Rhode Island, South Carolina, South Dakota, Vermont, Washington, District of Columbia, and Puerto Rico.

> to aid us in developing better evaluation reports. Given the tortoise and hare complex of USOE (in which it takes a year or two to get regulations devised and then allows no time at all to move into an implementation phase) it would appear that we cannot anticipate the new evaluation system being in place before FY 1979. (p.9)

Other states seem to be adopting the new reporting methods with more enthusiasm. For example, the Washington, D.C. report for FY 1976, adopted the NCE metric not only for reporting math and reading achievement test data for grades 1-3, but also for reporting kindergarten results on the Comprehensive Test of Basic skills, Form S (p. 52); self-concept scores from grades K-3 (pp. 83-86); and psychomotor outcomes from a random sample of first, second and third grades (pp. 103-105).

Approaches to evaluating outcomes of Title I early childhood programs seem to vary widely both across the PK, K and grade 1 levels and across the states. Table 4.7 summarizes information on which state reports indicate outcome evaluations at each of the early childhood levels. More than 40 states indicate that some sort of outcome evaluations were performed for first grade Title I participants. Around two dozen states indicated that such evaluations were carried out at the kindergarten level. At the PK level, however, only nine state reports gave indication of outcome evaluations.

Even when outcome evaluations were performed at the three early childhood levels, test results at these levels often were not reported. Reasons given for not reporting outcome data at the early childhood levels varied. Hawaii stated that kindergarten results were not reported because the set of tests used to evaluate kindergarten outcomes "was not compatible to reporting on a grade equivalent basis" (p. 31). In similar fashion, Maryland decided not to report kindergarten scores in grade equivalent form: "Because the scores of the majority of tests administered in kindergarten could not be converted to grade equivalent scores, the average or mean stanine score is provided for this summary of readiness scores" (p. 25). Curiously however, the grade

Table 4.7 SUMMARY OF REVIEW OF FY 1976 STATE TITLE I EVALUATION REPORTS: OUTCOME EVALUATIONS AT PK, K AND GI LEVELS

| State [1] | Indications of Outcome Evaluations [2] Pk | K | 1 |
|---|---|---|---|
| Alabama | x | x | x |
| Alaska | - | - | x |
| Arizona | - | - | - |
| Arkansas | - | - | x |
| California | - | - | x |
| Colorado | x | x | x |
| Connecticut | - | x | x |
| Delaware | - | x | x |
| Florida | - | x | x |
| Georgia | x | x | x |
| Hawaii | x | x | x |
| Idaho | - | - | x |
| Illinois | ? | ? | ? |
| Indiana | x | x | x |
| Iowa | - | - | x |
| Kansas | - | x | x |
| Kentucky | - | x | x |
| Louisiana | - | x | x |
| Maine | - | x | x |
| Maryland | - | x | x |
| Massachusetts | x | x | x |
| Michigan | - | - | x |
| Minnesota | x | x | x |
| Mississippi | - | - | - |
| Missouri | - | - | x |
| Montana | - | - | x |
| Nebraska | - | - | x |
| Nevada | - | x | x |
| New Hampshire | - | ? | - |
| New Jersey | - | - | x |
| New Mexico | - | x | x |
| New York | x | x | - |
| North Carolina | - | x | x |
| North Dakota | - | - | - |
| Ohio | - | x | x |
| Oklahoma | x | ? | x |
| Oregon | - | - | x |
| Pennsylvania | - | - | x |
| Rhode Island | - | x | x |
| South Carolina | - | - | - |
| South Dakota | - | - | - |
| Tennessee | - | - | x |
| Texas | - | - | - |
| Utah | - | - | - |
| Vermont | - | - | x |
| Virginia | - | - | x |
| Washington | - | - | x |
| West Virginia | - | - | x |
| Wisconsin | - | - | x |
| Wyoming | - | - | x |
| District of Columbia | - | x | x |
| American Samoa | | | |
| Guam | | | |
| Puerto Rico | - | x | x |
| Trust Territory | | | |
| Bureau of Indian Affairs | - | x | x |
| Virgin Islands | - | - | - |

Footnotes

1. There are a total of 57 "states" from which reports are potentially available. From three such units (American Samoa, Guam and Trust Territory), no reports were available. From five states (Gerogia, Massachusetts, Michigan, New Jersey, and West Virginia), only FY 1975 reports were available. Full citations for the reports are provided in appendix I.

2. X's indicate that some form of outcome evaluations with standardized or criterion-referenced tests are reported at the indicated levels. Dashes mean that there is no clear-cut indication of outcome evaluation. Question marks indicate that there is no clear-cut indication of outcome evaluations, but there is strong reason to presume that some kind of outcome evaluation was conducted.

equivalent issue did not stymie everyone. Oklahoma, for example, reported gains in grade equivalent units at the "pre-k" level as well as at higher grade levels.*

Iowa reported that no product evaluation data were included in the state report for either grade one or grades nine-twelve "due to the lack of subjects tested and the measuring instruments available" (p. 36). Evaluators in Missouri observed:

> There are problems...in establishing pretest data for first graders. Because of this problem first grade results were omitted for the state summaries except for summer school. (p. 19)

South Dakota did not report any outcome data for the early childhood levels, commenting:

> Testing data for kindergarten, first grade, and high school was inconsistent. It was either not available from the tests used, or if it was available the gains were often excessive. (p. 14)

Several states apparently did not consider "excessive" or unusually large gains at the early childhood levels as sufficient grounds for not reporting data. Arkansas, for example, reported Title I achievement gains in reading and math by grade level and found that the highest average gain in reading was for first graders--6 NCE's, or almost one-third of a standard deviation (p. 47). In math, Arkansas first graders made an average gain of eleven NCE's, exceeded only by fourth graders who gained fourteen NCE's on the average (p. 49).**

---

*The FY 1976 report from Oklahoma reports data from "pre-k" and first grade levels, but does not report kindergarten data separately. One suspects that the term "pre-k" may be used in this report to include both prekindergarten and kindergarten, but we have not checked this supposition with officials in Oklahoma.

**The Arkansas data are based on an unspecified sample of schools "with various program activities and using different types of standardized tests" (p. 44).

The issue of gains by grade-level was also explored in the District of Columbia's annual report. Evaluators there found that first grade gains averaged 4.7 NCE's in reading and 6.5 NCE's in math.* In an effort to obtain a clearer understanding of achievement differences between students in Title I and non-Title I schools, the Washington, D.C. evaluators analyzed data from two criterion-referenced tests which had been administered to first, second and third-graders. They found that "students in Title I schools show a clear advantage on several reading and mathematics objectives at grade one, but no such advantage is evident at grades two and three" (p. 64).

This tendency for greatest gains to show up in earlier intervention efforts harkens back to several points noted earlier. First, it seems to echo the theory regarding the plasticity of early development. Second, it also reflects the trend noted above for earlier interventions to show up frequently in prorams identified as examplary on the basis of impact data. Another explanation for this phenomenon is suggested by the District of Columbia evaluators. However, since this is closely tied to measurement issues, we will postpone elaboration on it until chapter V. For now, let us summarize not only what's been reviewed in this section, but more broadly what's been said in the whole of chapter IV.

---

*These estimates of program gains were based on data from the CTBS/Form S.

## 4.7 Summing Up

Early childhood programs have received only slim attention in past national evaluations of Title I. The only sort of study where early programs have shown up with any regular prominence is in efforts to identify exemplary programs. Why early interventions often show up as exemplary may be explained by a variety of factors; for example, because of special efforts in recent years to develop early programs; because of the presumed plasticity of early development; because of the relatively large scope of early interventions; or because of the means by which exemplary efforts have been identified. The last possibility will be explored more thoroughly in the next chapter.

We have relatively little information on the scope of early childhood Title I programs around the country. What information we have been able to piece together suggests that a substantial portion of state Title I allotments have been invested in early programs in at least twenty states.

Descriptive information on early childhood Title I programs is meager. What information we do have suggests that programs differ substantially at the prekindergarten, kindergarten, and first grade levels. At the prekindergarten level Title I often appears to fund entire intervention efforts; either home- or center-based. At the kindergarten level, Title I seems in several states to have served as a catalyst to promote state funding of expanded kindergarten programs. Of the three early childhood levels, most emphasis seems to have been placed on first-grade programs. First-grade Title I participants outnumber PK and K participants in all but one of the states from which disaggregated data are available. At the grade-one level, unlike the lower levels, Title I funds appear to be used mainly

to supplement ongoing educational efforts and to be aimed especially at bolstering school readiness skills. Also, programs at this level rarely seem to be focused solely on serving first-graders. More often they appear as broader efforts to aid educationally needy children in one-to-three or one-to-six grade spans.

Selection of children to participate in early programs poses a real problem, both because of measurement problems at the early levels and because uniform records are not available for children before they start their school careers. As a result, a wide variety of methods appear to be used to select participants for early childhood Title I efforts.

Title I rules and regulations call for parental involvement in Parent Advisory Councils at the school and district levels. Moreover, parent involvement seems especially important in early programs--especially in home-based intervention efforts. Despite its apparent importance, parental involvement seems to have received relatively little attention in early childhood evaluations, and to have widely varied support in different states around the country.

Evaluations of early childhood programs seem to present a host of special issues. Outcome evaluations are far more common at the first-grade level than at the K or PK levels. Even when tests have been administered at the early childhood level, data are often not reported for a variety of reasons. Many of the problems of evaluating the impact of early programs are connected to problems in early childhood testing and instrumentation. These will be directly addressed in chapter V.

Perhaps the best way to sum up the broader issues regarding evaluations of early childhood Title I programs is simply to quote a plaint expressed in

the FY 1976 Title I evaluation report from the State Department of Education in New Hampshire regarding early childhood efforts in that state.

> [Child Development Projects] seem to be a catchall category for those projects which are clearly working in the area of pre-school education and are very concerned with child development, those projects that do not seem to be designed to fit easily into an evaluation pattern using hard cognitive data and those who have used child development as a category in which they can perhaps use some sort of soft evaluation because they don't believe in standardized testing. The Department and the people in the field need to take a serious look at answering the questions "What is a child development project" and "How do you evaluate one?" (p. 16)

Until we have better answers to the former sort of question regarding what early childhood Title I programs are, we likely will not make much progress on the latter, how to evaluate them.

## V. EARLY CHILDHOOD TESTS AND INSTRUMENTS

As was evident in the last section of Chapter IV, some of the most prominent worries over evaluating Title I early childhood programs concern appropriate tests and measurement instruments. Hence, in this chapter we turn to consider this topic directly. The chapter is divided into the following sections:

- Instruments Used in Evaluations of Early Childhood Title I Programs
- Characteristics of Young Children that Make Testing Difficult
- Types of Instruments for Use With Young Children
- Technical Quality of These Instruments
- Use of Early Childhood Tests

### 5.1: Instruments Used in Evaluations of Title I Early Childhood Programs

From our analysis of FY 1976 State Title I Evaluation Reports, it is evident that a wide variety of test instruments have been used to evaluate Title I programs. Unfortunately, however, it is often impossible to tell which instruments were used because several states simply did not clearly identify which specific tests have been used. In many other cases, tests may be identified, but the grade levels at which they are used are not indicated. Even in cases where tests are identified by grade level, it is sometimes impossible to tell exactly which test is referred to because only brief incomplete test titles are indicated, such as the "California," or the "Metropolitan." Hence, in summarizing which tests were reported used in how many states at which level, we have simply had to make some guesses on the basis of brief names as to which tests were actually referred to. Nevertheless, Table 5.1 summarizes the tests used at first grade and PK-K or preschool level as far as can be determined from the FY 1976 state reports.

TABLE 5.1: Tests Reported Used in Title I Evaluations of the Early Childhood Level*

| Test Name | No. of States Reporting Use of Test by Level: PK-K or Preschool | First Grade |
|---|---|---|
| Boehm Test of Basic Concepts | 2 | 1 |
| Botel Reading Inventory | - | 2 |
| California Achievement Test | 1 | 11 |
| Comprehensive Test of Basic Skills | 3 | 8 |
| Diagnostic Reading Scale (Spache) | - | 3 |
| Durrell Analysis of Reading Difficulty | - | 4 |
| Gates-MacGinitie Reading Test | 3 | 17 |
| Gray Oral Reading Test | - | 3 |
| Holt Checklist (Diagnostic Test) | 1 | 1 |
| Iowa Test of Basic Skills | - | 12 |
| Key Math Diagnostic Arithmetic Test | - | 3 |
| Lee Clark Reading Readiness Test | 1 | 2 |
| Metropolitan Achievement Test | 2 | 12 |
| Metropolitan Readiness Test | 6 | - |
| Nelson Reading Test | - | 4 |
| Peabody Individual Achievement Test | - | 5 |
| Peabody Picture Vocabulary Test | 3 | 2 |
| Silveroli Reading Inventory | - | 3 |
| Slossen Oral Reading Test | 1 | 3 |
| SRA Achievement Series | - | 8 |
| Stanford Achievement Tests | - | 14 |
| Stanford Diagnostic Reading Test | - | 2 |
| Stanford Early School Achievement Test | 3 | 1 |
| Test of Basic Experiences | 5 | 1 |
| Wide Range Achievement Test | - | 13 |
| Woodcock Reading Mastery Test | - | 5 |

*This table lists only those tests reported used in two or more states at the early childhood level. In numerous cases, a test is reported used at a wider grade span than concerns us here; for example, at grades 1-3 or 1-6. In such cases, we assume that the test was in fact used at the first grade level.

This table can be compared with the findings of a previous study by Gamel et al, shown in Table 5.2. Apparently, the most widely used tests at the first grade level are part of series most commonly used generally; namely the California Achievement, the Gates-MacGinitie, the Iowa, the Metropolitan, the Stanford and the Wide Range series of tests. At the PK-K level, however, far fewer tests are reportedly used and the mix of tests reported used at this level is quite different than for later grades. More prominent are readiness tests, for example the Metropolitan Readiness Test. Also prominent at the PK-K level is the Test of Basic Experiences (TOBE), an instrument specifically designed to assess the "richness of conceptual background" of children in preschool through first grade (Buros, 1972).

What difference does it make what test is used to evaluate Title I interventions? Apparently, it can make quite a difference. This was shown in an unusual study reported by Wisconsin in that state's FY 1976 Title I evaluation report. The Wisconsin SEA performed an analysis to identify factors associated with student gain scores reported by LEAs. The analysis tested the statistical significance of the relationship between gain scores reported (in reading and in math) and the following variables:

1) number of months between pre-and post-test;
2) type of achievement test used;
3) students' grade placement; and,
4) number of years students had been in Title I.

For reading, the type of achievement test used was found to have a stronger association with amount of gain than any of the three other variables. In mathematics, interval between pre-and-post-tests was slightly more strongly associated with gains than type of test, but both were more strongly related to

TABLE 5.2

Frequency with which Various Standardized Tests were Cited in the State Title I Annual Evaluation Reports Presented by Fiscal Year *

| Standardized Test | Fiscal Year 69 | 70 | 71 | 72 | 73 | 74 |
|---|---|---|---|---|---|---|
| California Achievement Test | 29 | 20 | 17 | 25 | 23 | 7 |
| Comprehensive Test of Basic Skills | | | 5 | 10 | 9 | 2 |
| Durrell-Sullivan Test of Reading Capacity | 4 | 4 | 5 | 5 | 3 | 2 |
| Gates-MacGinitie Reading Test | 16 | 23 | 29 | 26 | 29 | 9 |
| Gray Oral Reading Test | 6 | 2 | 1 | 3 | 5 | 2 |
| Iowa Tests of Basic Skills | 21 | 19 | 16 | 19 | 19 | 6 |
| Lee-Clark Readiness Test | 3 | | 1 | 2 | 4 | |
| Metropolitan Achievement Test | 18 | 20 | 26 | 29 | 24 | 7 |
| Nelson Silent Reading Test | | | 2 | 5 | 1 | 1 |
| Peabody Picture Vocabulary | 2 | 2 | 1 | 6 | 10 | 6 |
| Sequential Test of Educational Progress | | | 1 | 1 | | |
| SRA Achievement Test | 8 | 8 | 17 | 16 | 14 | 8 |
| Stanford Achievement Test | 29 | 23 | 27 | 26 | 22 | 7 |
| Wechsler Intelligence Scale | 2 | | 1 | | 1 | |
| Wide Range Achievement Test | | | 6 | 10 | 17 | 9 |
| Other | 6 | 5 | 38 | 46 | 55 | 30 |
| Number of State Reports included | 46 | 45 | 50 | 52 | 52 | 30 |

*As pretests, posttests or both, at any grade level.

Source: Adapted from Gamel et al, State ESEA Title I Reports... (1975) p.29

gains than either grade placement or years in Title I. The report observes:

> Five of the major reading tests and two of the major mathematics tests showed a greater proportion of students having high rates of achievement than did the other major tests considered. (p.87)

Apparently, in some evaluations of Title I effectiveness, the measure may be the message.

No similar study has been performed, as far as we know, specifically for early childhood test results. Inspection of early childhood test gains in several state reports suggests, however, that for early learning, the measure may also sometimes be the message. For example, kindergarten average test score gains on the TOBE appear to be especially large in two state reports -- 35 to 40 percentile points on different subtests in Alabama, (p.46), and 34 percentile points in North Carolina (p.34). Also, first grade reading gains on the WRAT are reported to be especially large in several states, for example, 1.66 average months gained per month in Massachusetts (p.xx), and in West Virginia, 1.8 average months gained per month in the program (p. 10).

If the measure is the message in evaluating early childhood program evaluations, it is vital that we examine what it is that these measuring instruments are and how good they are. This we will do in later sections of this chapter. First, however, in section 5.2, we consider the characteristics of young children that complicate the use of test instruments with them.

## 5.2 Characteristics of Young Children that Make Testing Difficult

Before addressing technical issues concerning testing of young children, it is useful to note the characteristics of young children that often make traditional techniques of educational testing difficult or even impossible for use with them.

First of all, young children have little interest in test performance. They are much less motivated than older children by competition or the desire to measure up to standards (Terman and Merrill, 1973). They do not easily become absorbed in an externally imposed task and their attention is easily distracted (Pikunas, 1976). They are quickly fatigued (Illingsworth, 1974), and their responses may be influenced by hunger, restlessness, desire to please, timidity or fear of the test situation, and a multitude of other motives and circumstances.

In general, it is difficult to differentiate cognitive and affective aspects of young children's development (Bradley and Caldwell, 1974). Cognition and affect seem to develop together and to interact, making the measurement of one dependent upon the other until children are socialized into school and society and affective behavior becomes more stable. Related to this is the fact that young children simply do not have the self-regulatory abilities like those that allow older children to adapt their behavior to particular situations. For example, fear and anger are frequent reactions of four and five-year-olds appearing and disappearing rapidly, seemingly without warning (Pikunas, 1976). This reactivity makes young children's performance in testing and assessment situations particularly susceptible to extraneous considerations. For example, it has been shown that young children's test performance is much more influenced by situational variables--including the ethnicity of the test giver--than is the performance of older children (Epps, 1974).

Another issue is that most children in the prekindergarten to grade-one range cannot read. Even those who have attained some degree of reading proficiency may find it difficult to transfer this new acquisition to the test situation. Reading directions, finding and holding one correct test item, and then moving on to the next, transferring attention from the test booklet to a corresponding location or an answer sheet and actually writing, circling or filling in a correct response are themselves complex tasks that compete with the test's objectives. Anastasi (1976) observes that "In testing children below the fifth grade, the use of any separate answer sheet may significantly lower scores... at these [lower] grade levels, having the child mark the answers in the test booklet itself is generally preferable." (p. 34).

In sum, several characteristics of young children--their inability to read, short attention span, and susceptibility to extraneous influences in assessment situations--all make traditional standardized testing extremely difficult. These factors apparently have contributed as we shall see in following sections to the development of instruments for young children that more often than not have to be individually rather than group-administered and also have relatively low technical quality.

## 5.3 Types of Instruments

In light of the special problems associated with testing young children, what sorts of tests are available for use with them? In order to answer this question, we have reviewed and analyzed data from six major sources of information on educational tests, each published since 1970.* Each of these sources, together with the brief names by which we will call them hereafter are listed in Table 5.3. At the end of this section we will comment briefly on two reviews of observation systems which might be used in educational assessment.

---

*There are two other major compilations of information on tests: Goldman and Saunders (1974) and Chun, Cobb, and French (1975). However, neither of these documents describes instruments in terms of grade or age-level for which they are appropriate. Hence, we have not drawn on these sources in our review.

Table 5.3 Major Sources of Information on Tests

| Brief Name | Full Title |
|---|---|
| 7MMY | O. Buros, Seventh Mental Measurements Yearbook (Highland Park, N.J.: 1972) 2 volumes. |
| PSKTE | R.Hoepfner, et al; CSE-ECRC Preschool/Kindergarten Test Evaluations (Los Angeles, CA: UCLA Graduate School of Education, 1971). |
| ESTE | R. Hoepfner, et al. CSE Elementary School Test Evaluations (Los Angeles, CA: UCLA Graduate School of Education, 1976). |
| TMCDI | O.G. Johnson and J. Bommarito,Tests and Measurements in Child Development (San Francisco, CA: Jossey Bass, 1971). |
| TMCDII | O.G. Johnson,Tests and Measurements in Child Development Handbook II (San Francisco, CA: Jossey Bass, 1976) 2 volumes. |
| SEM | D.K. Walker, Socioemotional Measures for Preschool and Kindergarten Children (San Francisco, CA: Jossey Bass, 1973). |

Table 5.4 presents results of our search of Buros 7MMY for tests indicated as appropriate for the early childhood levels* The categories by which we have summarized this information are simply those employed by Buros himself. Many tests are listed in 7MMY as appropriate for more than one grade-level. Hence, the numbers of tests listed separately at the PK, K and G1 levels generally total more than the actual number of tests, listed as appropriate for the whole PK-G1 range.

Roughly one-third of the tests listed in 7MMY are appropriate for at least one of the early childhood levels (175 out of 511). Most of these (164) are appropriate for use at the first grade level. Substantially fewer (99) are intended for use at the kindergarten level, and still fewer (54) are indicated as appropriate at the prekindergarten level. The category in which most PK tests are available is individual intelligence tests. Also a fair number (10) of PK tests are listed in 7MMY in the sensory motor category. At the kindergarten level there are twenty-one readiness tests listed in 7MMY and only slightly fewer (18) individual intelligence tests. At the first grade level, the categories with the most tests listed are readiness (26), individual intelligence (21) group intelligence (19), sensory motor (17) and nonprojective character tests (17). Note that at the first grade level, tests categorized strictly as reading tests appear for the first time (12). Perhaps the most striking aspect of this table is that there are almost three times as many tests indicated as appropriate at the first grade level as are indicated appropriate at the PK level.

---

* Note that the Seventh Mental Measurements Yearbook cover only those tests published or revised since the Sixth Mental Measurements Yearbook in 1965. Also some tests published since 1965 are not included in 7MMY but are covered in Reading Tests and Reviews (Buros, 1968) or in Personality Tests and Reviews (Buros 1970).

Table 5.4 Early Childhood Tests and Instruments Listed in Buros Seventh Mental Measurements Yearbook (1972).*

| Category | Prekindergarten | Kindergarten | First Grade | Total PK-1 | Total All Ages |
|---|---|---|---|---|---|
| Achievement Batteries | 1 | 5 | 9 | 11 | 36 |
| Intelligence-Group | 3 | 12 | 19 | 19 | 74 |
| Intelligence-Individual | 18 | 18 | 21 | 24 | 33 |
| Intelligence-Specific | 3 | 5 | 5 | 5 | 14 |
| Mathematics | 1 | 2 | 7 | 7 | 95 |
| Miscellaneous | 1 | 1 | 1 | 1 | 10 |
| Multi-aptitude Batteries | - | 1 | 1 | 1 | ? |
| Reading | - | - | 12 | 12 | 30 |
| Diagnostic Reading | - | 1 | 7 | 7 | 13 |
| Miscellaneous Reading | - | - | 5 | 5 | 9 |
| Oral Reading | - | 1 | 3 | 3 | 3 |
| Readiness | 8 | 21 | 26 | 29 | 29 |
| Sensory Motor | 10 | 11 | 17 | 18 | 20 |
| Character Nonprojective | 4 | 9 | 17 | 19 | 123 |
| Character Projective | 5 | 12 | 14 | 14 | 22 |
| | 54 | 99 | 164 | 175 | 511 |

*Listings in Buros do not always indicate grade levels for which tests are appropriate. Hence in some cases we have used the following age-grade level conversions in summarizing the above information: prekindergarten=less than 4.5years of age; kindergarten=4.5to 6 years; first grade=6 to 7 years.

Table 5.5 shows the number of subtests listed in the CSE-ECRC Preschool/ Kindergarten Test Evaluations (PSKTE) as relevant to different educational domains and goals. In summarizing this information from PSKTE we have again simply relied upon the test classifications which this volume's authors employed. Unfortunately, the categories of Buros and the CSE-ECRC evaluators do not correspond very well, but still we can see the domains for which early childhood tests are most often available. Note also, that unlike Buros, the CSE-ECRC evaluators have focused upon subtests. This fact in part explains which PSKTE book contains so many more listings at the prekindergarten and kindergarten levels than does 7MMY. Nevertheless, the pattern shown in Table 5.5 is quite similar to that shown in Table 5.4. At the PK and K levels most instruments are available in the cognitive and reading readiness categories. In the readiness category there are substantially more subtests available at the kindergarten (142) than at the PK level (62).

Also shown in Table 5.5 is the relatively small number of group-administered tests at the preschool level. The vast majority of tests (154 or 72%) listed in PSKTE as appropriate for the preschool level are individually administered instruments. At the kindergarten level, a somewhat smaller proportion of available instruments are individually administered (217 or 51%). Such individually-administered instruments are prevalent at these levels because as pointed out in the last subsection, several characteristics of young children make standardized group testing with them extremely difficult.

The PSKTE does not include information on first grade tests. For this information from the CSE group we must turn to a later volume by this group--the CSE Elementary School Test Evaluations. This volume contains information on commercially published tests in the first-to-eight grade range. In Table 5.6 we

TABLE 5.5: PRE-SCHOOL AND KINDERGARTEN SUBTESTS BY DOMAIN AND TYPE OF INSTRUMENT

| DOMAIN AND GOALS | PRESCHOOL | | | | KINDERGARTEN | | | |
|---|---|---|---|---|---|---|---|---|
| | No. of Subtests | Rating Instrument | Individually Administered | Group Administered | No. of Subtests | Rating Instrument | Individually Administered | Group Administered |
| **AFFECTIVE DOMAIN** | | | | | | | | |
| 1. Development of Personality | 5 | 4 | 0 | 1 | 24 | 13 | 2 | 9 |
| 2. Development of Social Skills | 4 | 4 | 0 | 0 | 19 | 11 | 2 | 6 |
| 3. Development of Motivation for Learning | 1 | 0 | 0 | 1 | 4 | 2 | 0 | 2 |
| 4. Development of Aesthetic Appreciation | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Subtotal | 10 | 8 | 0 | 2 | 47 | 26 | 4 | 17 |
| **INTELLECTUAL DOMAIN** | | | | | | | | |
| 5. Cognitive Functioning | 27 | 0 | 27 | 0 | 57 | 0 | 43 | 14 |
| 6. Creativity | 11 | 1 | 9 | 1 | 19 | 1 | 10 | 8 |
| 7. Memory | 20 | 1 | 15 | 4 | 31 | 0 | 21 | 10 |
| Subtotal | 58 | 2 | 51 | 5 | 107 | 1 | 74 | 32 |
| **PSYCHOMOTOR DOMAIN** | | | | | | | | |
| 8. Physical Coordination | 36 | 8 | 21 | 7 | 48 | 8 | 25 | 15 |
| **SUBJECT ACHIEVEMENT** | | | | | | | | |
| 9. Arts and Crafts | 5 | 0 | 0 | 5 | 7 | 0 | 1 | 6 |
| 10. Foreign Language | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11. Function and Structure of the Human Body | 2 | 0 | 2 | 0 | 3 | 0 | 2 | 1 |
| 12. Health | 2 | 2 | 0 | 0 | 3 | 3 | 0 | 0 |
| 13. Mathematics | 8 | 1 | 6 | 1 | 19 | 1 | 9 | 9 |
| 14. Music | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15. Oral Language Skills | 30 | 4 | 26 | 0 | 40 | 4 | 35 | 1 |
| 16. Readiness Skills | 62 | 0 | 48 | 14 | 142 | 0 | 63 | 79 |
| 17. Reading and Writing | 0 | 0 | 0 | 0 | 4 | 0 | 3 | 1 |
| 18. Religion | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 19. Safety | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 20. Science | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| 21. Social Studies | 2 | 1 | 0 | 1 | 3 | 1 | 1 | 1 |
| Subtotal | 111 | 8 | 82 | 21 | 222 | 9 | 114 | 99 |
| Totals | 215 | 26 | 154 | 35 | 424 | 44 | 217 | 163 |

Source: Hoepfner, R., et al. CSE-ECRC Preschool/Kindergarten Test Evaluations. Los Angeles: UCLA Graduate School of Education, 1971.

Table 5.6: SUBTESTS FOR GRADES 1-4

| DOMAIN | GRADE 1 No. of Subtests | GRADE 2 No. of Subtests | GRADES 3-4 No. of Subtests |
|---|---|---|---|
| 1. Affective & Personality Traits | 443 | 434 | 633 |
| 2. Arts & Crafts | 24 | 25 | 28 |
| 3. Career Education | 25 | 26 | 68 |
| 4. Cognitive & Intellectual Skills | 366 | 304 | 438 |
| 5. Foreign Language | 13 | 29 | 33 |
| 6. Language Arts | 58 | 98 | 199 |
| 7. Mathematics | 100 | 134 | 286 |
| 8. Music | 0 | 1 | 61 |
| 9. Perceptual & Motor Skills | 376 | 286 | 309 |
| 10 Physical Education and Health | 18 | 23 | 33 |
| 11. Reading | 468 | 481 | 592 |
| 12. Religion & Ethics | 11 | 10 | 12 |
| 13. Science | 8 | 9 | 19 |
| 14. Social Studies | 18 | 21 | 46 |
| | 1928 | 1881 | 2757 |

Source: R. Hoepfner, et al. CSE Elementary School Test Evaluations Los Angeles: UCLA Graduate School of Education, 1976).

have summarized information from this volume on subtests appropriate in the grades 1-4 range. In summarizing this information we have again simply employed the categories used by the UCLA authors. As was evident in information from 7MMY, there is apparent in the ESTE, a tremendous proliteration of instruments at the first grade level. The ESTE lists almost 2000 subtests designed for use at the first grade level. Roughly the same number of subtests are available for use at the second grade level and almost 3000 subtests are available for use at grades 3-4. The category with the largest number of subtests at the first-grade level is reading (468); followed by affective and personality traits (443); perceptual and motor skills (376); and cognitive and intellectual skills (366). In contrast, the 1971 PSKTE listed only four reading subtests as appropriate for the kindergarten level.*

Buros' 7MMY and the UCLA test evaluation books, PSKTE and ESTE, cover only published tests and instruments. In order to consider instruments which are available in unpublished form we must look elsewhere. In Tables 5.7 and 5.8 we show information summarized from Johnson and Bommarito's TMCDI (1971) and Johnson's TMCDII (1973). These volumes were designed to "keep researchers, evaluators and clinicians up-to-date on the most recent developments in unpublished measures of child behavior." (TMCDII, p. ix). In summarizing information from these volumes we have again simply employed the classification schemes used by the authors. Unfortunately the classification scheme used in these volumes is quite different than that used by Buros in 7MMY or by Hoepfner et al in either PSKTE or ESTE. Indeed the classification scheme in TMCDII is slightly different than that used in TMCDI. TMCDI (or Handbook I as it is sometimes called)was designed to serve as a central source of information on unpublished tests for children

---

* Unlike the PSKTE, the ESTE does not indicate whether instruments are rating instruments, or individual or group tests.

Table 5.7: Unpublished Early Childhood Instruments (1971)

| Category | Number of Early Childhood Instruments: Prekindergarten | Kindergarten | 1st Grade | Total # of Instruments |
|---|---|---|---|---|
| Cognition | | | | |
| a. Intelligence & School Readiness | 9 | 9 | 11 | 21 |
| b. Language & Number Skills | 4 | 4 | 6 | 18 |
| c. Specific Achievements | 2 | 3 | 3 | 12 |
| d. Cognitive Style/Cognitive Processes | 2 | 2 | 4 | 20 |
| e. Miscellaneous | 1 | 2 | 3 | 8 |
| Personality & Emotional Characteristics | | | | |
| a. Personality - General | 6 | 10 | 12 | 23 |
| b. Personality Variables | 9 | 5 | 9 | 28 |
| c. Personality Adjustment | 1 | 4 | 6 | 15 |
| d. Anxiety | 1 | 0 | 5 | 7 |
| Perceptions of Environment | | | | |
| a. Attitudes Toward Adults | 6 | 4 | 5 | 19 |
| b. Attitudes Toward Peers | 2 | 1 | 1 | 7 |
| c. Other Factors | 0 | 1 | 2 | 8 |
| Self Concept | 1 | 2 | 3 | 14 |
| Motor Skills, Brain Injury, Self-Perception | | | | |
| a. Motor Skills | 2 | 3 | 3 | 8 |
| b. Brain Injury | 3 | 0 | 3 | 9 |
| c. Sensory Perception | 2 | 2 | 4 | 7 |
| Physical Attributes | 2 | 2 | 3 | 9 |
| Miscellaneous Attitudes & Interests | 1 | 3 | 2 | 12 |
| Social Behavior | 10 | 10 | 10 | 26 |
| Unclassified | 0 | 1 | 1 | 6 |
| Totals | 64 | 68 | 96 | 277 |

* Listings in TMCD I often do not indicate grade levels for which instruments are appropriate. So in this summary, we have used the following age-grade level conversions: 3-5 years = Prekindergarten; 5-6 years = Kindergarten; and 6-7 years = First Grade.

Source: Johnson, O. G. & Bommarito, J. Tests and Measurements in Child Development. (San Francisco: Jossey Bass, 1971).

Table 5.8: Unpublished Early Childhood Instruments (1976)

| Category | Numbers of Early Childhood Instruments | | | Total # Instruments |
|---|---|---|---|---|
| | Prekindergarten | Kindergarten | 1st Grade | |
| **Cognition** | | | | |
| a. Intelligence, Readiness Maturation | 46 | 51 | 47 | 92 |
| b. Language & Number Skills | 36 | 31 | 27 | 69 |
| c. Specific Achievements | 9 | 8 | 16 | 41 |
| d. Cognitive Style/Cognitive Processes | 18 | 26 | 31 | 63 |
| e. Miscellaneous | 1 | 1 | 2 | 4 |
| **Personality & Emotional Characteristics** | | | | |
| a. General Personality | 16 | 16 | 21 | 36 |
| b. Personality Variables | 10 | 13 | 17 | 61 |
| c. Personality Adjustment | 9 | 11 | 15 | 25 |
| d. Anxiety | 0 | 0 | 0 | 13 |
| **Perceptions of Environment** | | | | |
| a. Perceptions of Adults | 5 | 7 | 11 | 26 |
| b. Perceptions of Peers | 1 | 3 | 3 | 8 |
| c. Other Environmental Perceptions | 4 | 6 | 8 | 17 |
| **Self Concept** | 5 | 7 | 14 | 45 |
| **Motor Skills and Sensory Perception** | | | | |
| a. Motor Skills | 9 | 23 | 27 | 38 |
| b. Sensory Perception | 20 | 28 | 30 | 50 |
| **Physical Attributes** | 1 | 1 | 1 | 6 |
| **Attitudes & Interests** | | | | |
| a. Towards School | 1 | 7 | 15 | 46 |
| b. Miscellaneous | 8 | 8 | 15 | 71 |
| **Social Behavior** | 17 | 21 | 26 | 47 |
| **Vocational** | 0 | 0 | 0 | 17 |
| **Unclassified** | 6 | 6 | 10 | 27 |
| **Totals** | 222 | 274 | 336 | 802 |

* Listings in TMCD II often do not indicate grade levels for which tests are appropriate. So in summarizing information for this table we used the following age-grade conversions: 3-5 years = prekindergarten; 5-6 years = kindergarten; and 6-7 years = first grade.

Source: Johnson, O.G. Tests and Measurements in Child Development: Handbook II. (San Francisco: Jossey Bass, 1976) 2 Volumes.

aged 0-12, reported in the educational literature from 1956 to 1965. TMCDII or Handbook II, is essentially an extension of Handbook I covering unpublished measures reported in the literature from 1966 to 1974. The later volume was also an extension of the earlier volume in that it covers instruments for an expanded 0-18 year-old age range, and it is based on a more extensive literature search than was carried out for TMCD I.* The expanded age range probably led directly to inclusion of the new category "vocational" tests in the second volume.

Despite the varied labels and classification schemes used to organize information in these volumes, comparing TMCD I and II with the listings of published tests in 7MMY, PSKTE and ESTE, suggests at least in rough fashion the domains in which test development work seems to have been concentrated in the last few years. Note first that more work in recent years seems to have been going into developing tests for younger children. Comparing numbers of tests available at the three early childhood levels; it is apparent that the ratio of first grade to prekindergarten tests is much lower in TMCDI and II than in 7MMY.** Also, judging from the numbers of tests listed in TMCDII it appears that more work in recent years has been devoted to developing instruments concerned with sensory perception, cognitive style and cognitive processes. Also a substantial number of instruments listed in TMCDII, (relatively speaking far more than in 7MMY,) are noncognitive ones, most categorized by Johnson as dealing with personality and emotional characteristics.

---

* For a fuller description of the information base on which these volumes were based see TMCDII pp. 4-14.

**We should note that there are alternative explanations to the differing proportions of types of published and unpublished tests. It may be for example that fewer PK and K tests are published simply because there is a smaller market for these tests than for first grade tests.

Increased attention in the field of educational measurement to the non-cognitive attributes of young children is also evident in a recent book on exactly this topic--D.K. Walker's Sociemotional Measures for Preschool and Kindergarten Children (1973). Table 5.9 shows the numbers of socioemotional measures covered in Walker's book by the categories she used. Walker's categorization scheme is again somewhat different than those employed by others.

Perhaps the most striking feature of Walker's volume is that she discovered a total of 143 socioemotional measures designed for use with prekindergarten and kindergarten children. Less than half this number of noncognitive instruments for PK-K children were covered in Buros' 7MMY published just one year earlier (see Table 5.4). In part the greater number of noncognitive instruments covered in Walker may be attributed to her inclusion of unpublished instruments as well as published ones. Also, however, it seems to reflect a real increase in interest in the measurement of young children's non-cognitive development.

Instruments covered in the sources described above include not only traditional paper-and-pencil tests but also checklists, rating scales, and observation schedules. Several documents in recent years have focused more narrowly on the latter type of instruments, that is observation systems. For example, A. Simon and E.G. Boyer's Mirrors of Behavior (1967, 1970) reviewed 79 different observation systems. Mirrors of Behavior does not indicate age-or grade-levels for which the systems are appropriate. Nevertheless, brief inspection of these volumes suggests that very few of the observation schemes were appropriate for or had been used with prekindergarten or kindergarten children. Since many of the instruments were developed for elementary classroom use it seems likely that they might be more appropriate for use at the first-grade level than at earlier levels.

Table 5.9: Socioemotional Measures for Preschool and Kindergarten Children (1973)

| Category | Number of Tests: Preschool (3-5 years) | Number of Tests: Kindergarten (5-6 years) | Total # Available |
|---|---|---|---|
| Attitudes | 8 | 10 | 11 |
| General Personality and Emotional Adjustment | 22 | 36 | 38 |
| Interests or Preferences | 5 | 5 | 7 |
| Personality or Behavior Traits | 34 | 22 | 41 |
| Self-Concept | 11 | 14 | 18 |
| Social Skills or Competency | 23 | 24 | 28 |
| | 103 | 111 | 143 |

* Listings in SEM often do not indicate grade levels for which instruments are appropriate; so in summarizing information from this volume, the following age-grade conversions were used: 3-5 years=prekindergarten; and 5-6 years= kindergarten.

Source: D.K. Walker, Socioemeotional Measures for Preschool and Kindergarten Children. (San Francisco: Jossey-Bass, 1973.)

Measures of Maturation: An Anthology of Early Childhood Observation Instruments (E.G. Boyer, A. Simon and G.R. Karafin, Eds.), 3 Volumes, Phila: Research for Better Schools, 1973) does describe the 73 observational systems reviewed according to age level. Of the 73 systems reviewed for the 0-6 age range, 51 (70%) reportedly can be used with preschoolers, 31 (42%) with kindergarteners, and 22 (30%) with first graders.

We will not here go into any detailed description of observation instruments save to note one especially important feature of some of them. Unlike most traditional paper-and-pencil tests which focus on individual children as the object of assessment, many of the observation systems focus on the classroom group or instructional setting as the object of evaluation. Instead of trying directly to assess educational outcomes for individual children such techniques focus on educational processes on the assumption that certain types of educational processes will have beneficial impact on children. This approach may be especially appropriate for evaluating early childhood interventions since, as we shall see in the next section there are some substantial technical problems involved in direct evaluation of educational outcomes for young children.

## 5.4 Technical Quality

Strictly speaking, the quality or worth of any assessment procedure resides not in the assessment instrument itself, but in the use of the instrument in a particular way with a particular population. As Cronbach observed, "one validates not a test, but an interpretation of data arising from a specified procedure." (Cronbach, 1971, p. 447). Standards for conducting such validation and assessing the technical quality of test instruments are not, however, altogether agreed upon. Regarding validity, for

example, Messick (1975) has argued for the primacy of construct validity over content or criterion-related validity. In an article entitled, "Must All Tests Be Valid," Ebel maintained several years ago (1961) that the meangingfulness of test scores may depend more on operational definition of measurement procedures themselves than on validation in terms of criterion variables on abstract constructs. Similar arguments have centered on standards for assessing test reliability. We cannot delve into such theoretical arguments here, but point them out simply to make clear that appropriate standards for judging technical qualities of tests are a topic of some debate.* Nevertheless, recognizing that technical quality of tests does not reside in the instruments themselves but in their use in particular ways, let us try to sum up what can be said about the technical qualities of tests for use with young children.

Unfortunately, most of the sources of information on instruments for use with young children provide little consistent data on the technical quality of those instruments for particular uses. In what is doubtlessly the best known source of information on published tests, Buros' 7MMY, reviewers offer their professional judgments on the quality of particular instruments. Nevertheless, the standards by which reviewers judge tests seem to vary markedly. Reviewers inclined toward the psychometric tradition seem to give more weight to considerations of reliability and validity, while clinically-oriented reviewers tend to place more credence in the apparent usefulness of an instrument in real - life settings, for example, in diagnosing learning problems. The only overall summary statement on the quality of tests covered in Buros' 7MMY, is Buros' own comment that "at least half of the tests currently on the market should never have been published" (p. xxviii).

---

* Even the official Standards for Educational and Psychological Tests of the American Psychological Association, the American Educational Research Association, and the National Council on Measurement in Education are now debating standards for judging technical qualities of tests.

The volumes by Johnson and Bommarito provide even less critical assessment of tests' quality than the 7MMY. For literally dozens of tests listed in TMCDI and TMCDII no data at all are reported pertaining to test validity and reliability. For other tests essentially meaningless information is provided, for example of the sort "reliability and validity were ascertained when the test was developed." In TMCDII, the only summary statement offered concerning tests' validity and reliability is Johnson's observation that "there are wide variations among the measures in the amount of supportive data collected and presented." (p.8).

Walker uses essentially the same format for describing tests as used by Johnson and Bommarito. For many of the 143 tests described, she notes that information on test norms, validity and reliability is simply not available. Unlike Johnson and Bommarito, however, she offers a more critical assessment of the overall technical quality of the tests she reviews:

> Very few instruments have adequate standardization norms that are representative for a wide range of children of varying ethnic groups, intelligence levels and socioeconomic backgrounds. Generally the ones that do exist are very poor and inadequate, since they are based on extremely small or narrowly defined populations of children. Many measures have minimal reliability data available for review. Reliability coefficients for those tests with adequate information are generally in the moderate range--from .50 to .80. Total scores are consistently more reliable than subpart scores or individual items. Most tests do not give enough information on reliability to judge whether the test is reliable for a variety of groups of children. The results of studies are generally based on very small samples that are not representative of the total population or even of one subgroup of preschool and kindergarten children. Often the ones with high internal consistency test-retest, or interrater reliability coefficients turn out to be inadequate measures because of their poor validity (p. 37).

Easily the most comprehensive review of the technical and practical qualities of tests for young children is provided by Hoepfner et al in their series of test evaluation reports. The UCLA investigators devised what they called the MEAN test evaluation system. The acronym stands for

> Measurement Validities,
> Examinee Appropriateness,
> Administrative Usability, and
> Normed Technical Excellence.

For each of these categories, Hoepfner et al devised a system whereby each test under review could receive from 0 to 15 points:

> The several categories of evaluation were assigned to those members of the UCLA team with appropriate expertise in that area. Each test was independently rated according to the MEAN system by at least two raters working without access to the other raters. For each subscale to be evaluated, each reviewer used a standard rating procedure...
>
> The primary concerns in the application of the MEAN system were the objectivity and consistency of the evaluations. To maximize both the objectivity for any one test evaluator and the consistency across evaluators, specific guidelines for evaluation of each aspect of each criterion and the following lettergrade assignment were adopted.
>
> Within any one of the four evaluative categories, a letter grade, based on the points assigned to each aspect of each criterion, was computed. The sum of the points, ranging from 0-15, was the determinant of the letter grades in the following manner:
>
> "G" (Good; 11-15 points). The test meets the criterion very well. Little improvement appears necessary for its immediate utilitization. CSE and ECRC [Center for the Study of Evaluation and Early Childhood Research Center] would endorse such a measure or employ it in their own assessment efforts.
>
> "F" (Fair; 6-10 points). The instrument is probably among the better tests available, but it does not completely meet the criteria. Alterations would probably render the test a candidate for "G" status. If no better qualified measure were available the Centers would employ an "F" measure, but would interpret results more cautiously.

"P" (Poor; 0-5 points). The instrument does not meet the criterion; it is clearly unsatisfactory in that area. Rather than employ such a measure for assessment, it would be better to seek an alternate device or attempt to develop one.

Each test then earned four letter grades by the MEAN system. The four letter combination serves as the Center's official evaluation of the test.*

This description of the test evaluation procedure employed by CSE comes from PSKTE. Basically, the same procedure was reported used in ESTE, except that in the latter volume somewhat more stringent criteria were applied. In PSKTE, a good rating was given when a subtest received 11-15 points in the relevant category; a fair rating when a subtest had 6-10 points; and poor when 0-5 points were received. However, in ESTE, the good rating was given only when a subtest received 12-15 points; fair for 8-11 points, and poor for 0-7 points.

Before reporting the results of CSE test evaluation results, we should note that the CSE effort has been sharply criticized by some measurement specialists. Anastasi (1978), for example, criticized the procedures used by the CSE test evaluators as "incredibly subjective". She suggested that the personnel who performed the test evaluations may not have been adequately trained and that some of the tests were inappropriately categorized by the CSE test evaluators. Similar issues were raised in a recent study by the General Accounting Office (GAO) entitled Problems and Needed Improvements in Evaluating Office of Education Programs (1977). The GAO commented on the

---

* Hoepfner et al, PSKTE, pp. xx-xxii. For details on how points were assigned within each of the four major categories, see PSKTE & ESTE.

CSE test evaluations as follows:

> There was general agreement among the educators and test publishers interviewed that the Center's work was a definite improvement over other existing evaluations of standardized tests. Some test publishers, however, criticized the Center's work because it was based on incomplete data, excluding, for example, the highly technical specifications from which test questions are developed. They also criticized it for establishing educational goals on which test ratings were based without consulting publishers, and for relying on the work of graduate students.
>
> In response, Center officials indicated that they had sent letters to test publishers advising them of the study and its purpose, stating what information was desired, and asking for any other material the test publishers wished to send. They acknowledged that their curricular goal system is not perfect; however, they believe it is a tremendous step forward, providing a clear statement of expected student behaviors which test users can employ to match tests to curriculums. Officials also stated that the system is justified because test makers often measure skills that are different from what they claim to measure. Center officials also noted that all participating graduate students demonstrate adequate competence in research and measurement before being hired; they followed set procedures, were routinely monitored, and discussed questionable points with supervisors. State education agency officials and measurement experts interviewed stated that the Center's work should be continually updated and should include an assessment of available criterion-referenced tests (pp. 47-8).

Hence, the reader should beware of the fact that there is some dispute over the manner in which the CSE-UCLA test evaluations were carried out. Nevertheless, since their work does seem to be "a definite improvement over other existing evaluations of standardized tests", it is worth close scrutiny.

How good are early childhood tests according to the UCLA ratings? Pertinent information on tests in the preschool to grade 4 range is shown in Tables 5.10 and 5.11. The second column under each grade level shows the number of tests, by goal area which are rated at least fair in each of the four categories in which tests were rated; that is, measurement validity, examinee appropriateness, administrative usability and normed technical excellence. In other words, excluded from the count of at least FFFF tests are all ones which the UCLA investigators deemed unsatisfactory in any of the four areas.

Table 5.10: MEAN Evaluations of Pre-School and Kindergarten Subtests

| DOMAIN AND GOALS | | PREKINDERGARTEN No. of Subtests | PREKINDERGARTEN No. Rated at Least FFFF | KINDERGARTEN No. of Subtests | KINDERGARTEN No. Rated at Least FFFF |
|---|---|---|---|---|---|
| AFFECTIVE DOMAIN | | | | | |
| 1. Development of Personality | | 5 | 0 | 24 | 0 |
| 2. Development of Social Skills | | 4 | 0 | 19 | 0 |
| 3. Development of Motivation for Learning | | 1 | 0 | 4 | 0 |
| 4. Development of Aesthetic Appreciation | | 0 | 0 | 0 | 0 |
| | Subtotals | 10 | 0 (0%) | 47 | 0 (0%) |
| INTELLECTUAL DOMAIN | | | | | |
| 5. Cognitive Functioning | | 27 | 0 | 57 | 10 |
| 6. Creativity | | 11 | 0 | 19 | 0 |
| 7. Memory | | 20 | 0 | 31 | 0 |
| | Subtotals | 58 | 0 (0%) | 107 | 10 (9%) |
| PSYCHOMOTOR DOMAIN | | | | | |
| 8. Physical Coordination | | 36 | 0 | 48 | 2 (4%) |
| SUBJECT ACHIEVEMENT | | | | | |
| 9. Arts & Crafts | | 5 | 0 | 7 | 0 |
| 10. Foreign Language | | 0 | 0 | 0 | 0 |
| 11. Function and Structure of the Human Body | | 2 | 0 | 3 | 0 |
| 12. Health | | 2 | 0 | 3 | 0 |
| 13. Mathematics | | 8 | 0 | 19 | 2 |
| 14. Music | | 0 | 0 | 0 | 0 |
| 15. Oral Language Skills | | 30 | 2 | 40 | 4 |
| 16. Readiness Skills | | 62 | 3 | 142 | 17 |
| 17. Reading and Writing | | 0 | 0 | 4 | 0 |
| 18. Religion | | 0 | 0 | 0 | 0 |
| 19. Safety | | 0 | 0 | 0 | 0 |
| 20. Science | | 0 | 0 | 1 | 0 |
| 21. Social Studies | | 2 | 1 | 3 | 1 |
| | Subtotals | 111 | 6 (5%) | 222 | 24 (9%) |
| | Total = | 215 | 6 (3%) | 424 | 36 (8%) |

Source: Ralph Hoepfner, et al. CSE-ECRC Preschool/Kindergarten Test Evaluations. (Los Angeles: UCLA Graduate School of Education, 1971.

Table 5.11: MEAN EVALUATIONS OF SUBTESTS FOR GRADES 1 - 4

| DOMAIN | GRADE 1 | | GRADE 2 | | GRADES 3-4 | |
|---|---|---|---|---|---|---|
| | No. of Subtests | No. Rated at Least FFFF | No. of Subtests | No. Rated at Least FFFF | No. of Subtests | No. Rated at Least FFFF |
| 1. Affective & Personality Traits | 443 | 0 | 434 | 0 | 633 | 0 |
| 2. Arts & Crafts | 24 | 0 | 25 | 0 | 28 | 0 |
| 3. Career Education | 25 | 0 | 26 | 0 | 68 | 0 |
| 4. Cognitive & Intellectual Skills | 366 | 6 | 304 | 3 | 438 | 9 |
| 5. Foreign Language | 13 | 0 | 29 | 0 | 33 | 0 |
| 6. Language Arts | 58 | 0 | 98 | 7 | 199 | 22 |
| 7. Mathematics | 100 | 1 | 134 | 6 | 286 | 20 |
| 8. Music | 0 | 0 | 1 | 0 | 61 | 1 |
| 9. Perceptual & Motor Skills | 376 | 0 | 286 | 0 | 309 | 0 |
| 10 Physical Education and Health | 18 | 0 | 23 | 0 | 33 | 0 |
| 11. Reading | 468 | 1 | 481 | 5 | 592 | 27 |
| 12. Religion & Ethics | 11 | 0 | 10 | 0 | 12 | 0 |
| 13. Science | 8 | 1 | 9 | 0 | 19 | 1 |
| 14. Social Studies | 18 | 0 | 21 | 0 | 46 | 0 |
| Total | 1928 | 9 (0%) | 1881 | 21 (1%) | 2757 | 80 (3%) |

Source: R. Hoepfner, et al. CSE Elementary School Test Evaluations. (Los Angeles: UCLA Graduate School of Education, 1976).

It is apparent from Tables 5.10 and 5.11 that as rated by the CSE investigators, early childhood tests are to a very large extent unsatisfactory. At the prekindergarten level, only 6 out of 215 subtest were rated at least FFFF. At the kindergarten level, only 36 of more than 400 subtests achieved this rating. At the first grade level, only 9 out of nearly 2,000 subtests were rated FFFF or better.* More minimally adequate tests were available at higher grade levels, but even there relatively few tests were rated as FFFF or better.

The only goal area in which as many as a half-dozen subtests were rated as at least FFFF, were areas of cognitive functioning at the kindergarten and first grade levels and readiness skills at the kindergarten level. The vast majority of tests in all other goal areas were rated as poor or unsatisfactory in one or more of the four evaluative categories rated by the UCLA investigators.

The evaluative categories on which most subtests received lowest ratings were measurement validity and especially normed technical excellence. Numerous tests received fair or good ratings on the examinee appropriateness and administrative usability categories but were rated poor in terms of measurement validity or normed technical excellence. Also, there was a tendency for subtests to be rated lower in terms of normed technical excellence than total test scores. For example, consider the rating given in PSKTE to the subtests of the Metropolitan Readiness Test for use at the kindergarten level:

| | | |
|---|---|---|
| Alphabet | - | FGGF |
| Copying | - | GGFP |
| Listening | - | FGGP |
| Matching | - | FGGP |
| Numbers | - | GFFP |
| Word Meaning | - | GGGP |
| Total | - | FFFF |

---

*We should note that the apparent diminution of test quality from the PK-K levels to the first grade level may be an artifact of the more stringent standards applied in ESTE as compared to those used in PSKTE.

Recall that Walker made precisely the same sort of observation with respect to socioemotional measures: "Total scores are consistently more reliable than subpart scores or individual items" (p.37).

These observations are not terribly new. In fact they reflect the very history of educational and psychological testing. In the 19th century, researchers like Cattell and Whipple attempted to measure specific mental attributes with a wide range of separate instruments. Their efforts failed, however, to produce any instruments which reliably could be used to predict academic performance. The breakthrough in mental testing came around the turn of the century with Alfred Binet's work in France. Binet abandoned the attempt to isolate specific mental faculties and instead grouped together in one test a range of tasks: memory problems, copying figures and word definitions for example. His work led, of course, to the general intelligence test, and to modern day itelligence tests. The Stanford-Binet Intelligence Scale received only PFPP ratings at both the PK and K levels in the PSKTE, but other intelligence tests, like the Wechsler intelligence scales received among the highest ratings of any of the early childhood scales rated by the UCLA investigators.*

This phenomenon presents a real dilemma for evaluators of early childhood intervention programs. The more specific skill outcomes for which an evaluator hopes to find a measuring instrument, the more likely it is such an instrument will be rated low in terms of normed technical excellence, especially in terms of reliability. Conversely, the more weight

---

* For example, the Wechsler Preschool and Primary Scale of Intelligence Vocabulary Scale was rated FGFF at both PK and K levels by the UCLA group (PSKTE, pp. P10, K15.)

evaluators give to technical considerations of reliability in selecting test instruments, the more they will tend to be pushed toward general intelligence type tests.

One might hope that new and better instruments will soon be developed to assess specific outcomes for young children--especially in the noncognitive realm which seems of special importance for young children. As the UCLA investigators suggested with regard to instruments which they rated poor: "Rather than employ such measure for assessment, it would be better to seek an alternative device or to develop one" (PSKTE, p. xxii). Instrument development is, however, a job that most test users do not have the time or resources to undertake. Also, this sort of advice must be judged in the light of previous experience. For example, in the national Follow Through evaluation vast effort went into developing a variety of new instruments for use in evaluating the Follow Through program. The effort to develop new instruments for the Follow Through evaluation, despite years of effort and hundreds of thousands of dollars of funding, failed almost completely. In retrospect Follow Through evaluators warned:

> The temptation in undertaking an evaluation is to assume that procedures can be rapidly developed or that ways to deal with problems of measurement and analysis will emerge. There is, however, great danger in basing one's plans on the assumption that such developments will occur...While the importance of measures in the noncognitive domain was clearly recognized, neither limitations in existing instruments...nor the difficulties inherent in effecting new developments were equally clearly recognized, and certainly not on a schedule that could be useful to the evaluation. Evaluation planning, in short, should be limited to the state of the art as known or best estimated at the time the evaluation begins so that unrealistic expectations about new developments can be avoided. (Emrick, Sorenson and Stearns, 1973, p. 323). *

---

*It should be noted that the instrument development effort in Project Follow Through was troubled by far more than just technical problems. For an account of why this effort went awry, see Haney, 1977.

## 5.5 Test Use

If evaluation instruments for use with young children are of such poor quality, how is it, as we saw in Chapter IV, that a good number of early childhood programs have been identified as exemplary -- that is, have been shown to produce measureable impact on some sorts of test instruments? The answer seems to lie in three interrelated considerations having to do with: (1) the type of instruments generally used to show the impact of early childhood programs; (2) the nature of early interventions; and, (3) the norms for early childhood tests. After discussing these points, we will turn in Chapter VI to the broader and more fundamental issue of whether the type of early childhood tests that have been used to demonstrate early childhood program effectiveness do, in fact, seem to provide the sort of information needed by decision-makers at the local, state and federal level.

Early childhood tests used in the past to demonstrate program effectiveness represent only a small range of the sorts of tests available. For example, in his wide ranging review of the effectiveness of early interventions, Urie Bronfenbrenner relied mainly on IQ scores and to a lesser extent on standardized measures of academic achievement as his criteria of program effectiveness (Bronfenbrenner, 1974, p. 2). This same pattern of test use was found in a recent study by Heyneman and Mintz (1976). These investigators carried out a study to determine whether or not there is any relationship between the frequency with which tests are used in federally-sponsored research on children and adolescents and the quality of those tests as rated by the UCLA test evaluators.* To assess frequency of test use, Heyneman and Mintz surveyed 3,538 project proposals funded by 23 federal agencies in FY 1975.

---

* Unfortunately, Heyneman and Mintz included in their measures of tests only the UCLA ratings pertaining to measurement validity and normed technical excellence. They excluded from their study UCLA ratings pertaining to examinee appropriateness and administrative usability.

From these proposals, they identified 1,570 instruments. They found that only 484 could be identified in standard test reference bibliographies. They discovered that intelligence tests and academic achievement batteries were used far more often than other types of tests:

> The few academic achievement batteries (5% of the titles) accounted for over a fourth of all tests and measures used (26%) and intelligence test titles (13% of total number of titles) accounted for almost a third of test usage. (pp. 3-4).

Unfortunately, Heyneman and Mintz did not report on frequency of tests used by grade level, and we know of no other source reporting specifically on the frequency of tests used at the early childhood level. An indirect source of information on this topic, however, is tests used to show effectiveness of Title I exemplary programs. We do not have comprehensive information on the sorts of instruments used to identify the exemplary programs described in Chapter IV, but from the information available, it seems that exactly these same sorts of tests have served as the primary criterion of effectiveness for these programs.* Specifically, the instruments used to identify exemplary or effective programs in the PK to grade 1 range seem mainly to fall within the two domain areas noted in section 5.4 as having relatively more tests of acceptable technical quality -- that is, tests of cognition and academic readiness skills.

Since demonstrable impact of early childhood interventions seem to have been limited mainly to these two types, a relevant question is the relationship between tests in these two areas. A close look at classifications of subtests

---

*For example, the following exemplary ECTI programs used the following instruments to demonstrate program effects;

- Title I ESEA Preschool Program, Bessemer, AL - TOBE
- Baptist Hill Kindergarten, Greenville, AL - TOBE
- Child Parent Centers Activity, Chicago, IL - Caldwell PSI
- PK Education for the Disadvantaged Child, Cedar Rapids, IA
  - Slossen Intelligence Test for Children
  - Boehm Test of Basic Concepts
  - ITPA Grammatic Closure Subtest

made by the UCLA investigators in PSKTE and ESTE suggests that there is considerable confusion as to the differences between "cognitive" tests and "readiness" tests. As is shown in Table 5.12, subtests of some intelligence tests were classified by the UCLA evaluators as "readiness skills" tests and several subtests of one readiness test were classified in the cognitive functioning area. One dramatic example of the confusion between these two types of tests is the Information and Comprehension subtest of the Pictoral Test of Intelligence. In 1971, in PSKTE, this subtest was classified as a readiness skills test at both the preschool and kindergarten levels. Yet, in 1976 in ESTE, this same subtest was classified in the cognitive domain as a test of understanding and reasoning at the first grade level! The ambiguity over the relationship between readiness tests and cognitive tests was also noted in the FY 1976 Title I report from New York:

> Districts developed objectives that sought improved prereading or premathematics skills for pupils of the prekindergarten and/or kindergarten age span. Measurement devices, developed to assess behavior in the cognitive area for this age span (4-6) often measure reading readiness or mathematics readiness skills simultaneously (e.g., left-to-right perceptions are necessary for following English language print and the initial steps of understanding the number line continuum). The outcome behaviors therefore, are not discrete to one subject matter area and are reported as general cognitively oriented readiness achievement. (p.8)

In short, even though tests used in identifying effective PK-Grade 1 programs may have been labeled differently -- as readiness, IQ, or basic concepts tests -- the real distinctions between these types of tests seem very unclear. At least some instruments that are labelled as "readiness tests" seem to cover the same sort of material as traditional intelligence tests or tests of general cognition.*

*This observation also, of course, raises some serious questions as to the validity of labels applied to different tests, but that issue is beyond the scope of our immediate interest here.

Table 5.12: Classification of Selected Subtests as Cognitive or Readiness Instruments in CSE-ECRC Test Evaluations

| PSKTE | Cognitive Functioning (5) | Readiness Skills (16) |
|---|---|---|
| Preschool (Prekindergarten (pp. P-2-3, P-11-13) | Illinois Test of Psycholinguistic Abilities<br>- Visual Association | Illinois Test of Psycholinguistic Abilities<br>- Auditory Closure<br>- Auditory Reception<br>- Composite<br>- Sound Blending<br>- Visual Closure |
| Kindergarten (pp. K-4-6, K-16-22) | Minnesota Preschool Scale - Nonverbal | Minnesota Preschool Scale - Total |
| | Parent Readiness Evaluation of Preschoolers<br>- Combined Total<br>- Concepts<br>- Opposites<br>- Performance Total<br>- Verbal Association<br>- Visual Interpretation<br>- Visual Motor Association<br>- Auditory Association | Parent Readiness Evaluation of Preschoolers<br>- Comprehension<br>- Identification |
| | Pictoral Test of Intelligence - Similarities | Pictoral Test of Intelligence<br>- Form, Discrimination<br>- Information and Comprehension<br>- Picture Vocabulary<br>- Total |
| | Wechsler Preschool & Primary Scale of Intelligence<br>- Block Design<br>- Comprehension<br>- Full Scale<br>- Full Scale Score<br>- Mazes<br>- Performance<br>- Similarities | Wechsler Preschool & Primary Scale of Intelligence<br>- Picture Completion |

| ESTE | Understanding & Reasoning (8) | Reading Readiness Skills (28) |
|---|---|---|
| Grade 1 (pp. 32-46, 88-106) | Illinois Test of Psycholinguistic Abilities<br>- Auditory Association<br>- Auditory Reception<br>- Composite<br>- Manual Expression<br>- Verbal Expression<br>- Visual Association<br>- Visual Reception | Illinois Test of Psycholinguistic Abilities (Revised)<br>- Auditory Closure<br>- Grammatic Closure<br>- Sound Blending |
| | Pictoral Test of Intelligence<br>- Information & Comprehension<br>- Picture Vocabulary<br>- Similarities<br>- Total Raw Score | Pictoral Test of Intelligence - NONE |
| | Wechsler Preschool & Primary Scale of Intelligence<br>- Full Scale IQ<br>- Block Design<br>- Mazes<br>- Picture Completion<br>- Verbal IQ<br>- Comprehension<br>- Similarities<br>- Vocabulary | Wechsler Preschool & Primary Scale of Intelligence - NONE |

These observations still do not help us answer the question of why early childhood programs seem to have been identified as exemplary in relatively high proportion as compared with later grades interventions. After all, later-grades programs could just as easily as early childhood ones rely on cognitive instruments as a means for demonstrating their impact. The difference seems to lie, as was suggested in section 4.7, in the magnitude of early interventions as compared with later grades programs. As evidence in section 4.3 indicated earlier Title I interventions seem to be more comprehensive, to amount to sharper interventions into the status quo than later grades interventions. And, as Urie Bronfenbrenner has pointed out, when children first enter school, they consistently seem to show a rise in IQ test scores.

> When the disadvantaged child receives additional cognitive stimulation, as he typically does upon entry into school, his capacity to perform on tests of cognitive function is enhanced. This is particularly true for a youngster who is exposed to an educational program for the first time...
> (Bronfenbrenner, 1974, p.18).

The effect of first schooling on children's test performance shows up quite clearly in the norms for early childhood tests - what we alluded to in section 4.6. The phenomenon to which we refer is that, as shown in early childhood test norms, young children's achievement improves much more dramatically in kindergarten and first grade than in later grades. The issue was treated in the FY 1976 Title I report from Washington, D. C. Evaluators there addressed the issue by asking how much growth does the average student make in reading achievement during first grade, and is it the same as the growth realized by average students in later grades? (p. 72). An answer to the question can be derived by examining how well a 50th percentile student would perform at beginning and end of each grade level. For example, at the beginning of first grade, the 50th percentile score on the Comprehensive Test of Basic Skills (Level B) Total Reading equals a raw score of 31. On the end of first grade norms, the same raw score is

comparable to later grades interventions. For if natural growth of children in the early grades is much larger than that of children in later grades, gains of one-third of a standard deviation or seven NCEs (often used as criteria of educational significance) may simply not mean the same thing at the PK, K or grade 1 levels as for later grades. Also, the phenomenon of large effects of first schooling suggests that evaluation of early childhood Title I programs at whichever level - PK, K or grade 1 - should not ignore the issue of what sort of schooling experience children have had prior to their participation in early childhood Title I programs.

In closing this section, we should note that this discussion of test use has focused mainly on one particular sort of test use -- namely the use of tests as criteria of program effectiveness. We have focused on this sort of test use mainly because this is the manner in which tests generally have been used in educational program evaluation. Nevertheless, in practice, tests also are likely used just as frequently for needs assessment as for program evaluation. As noted in section 4.4, the use of tests in early childhood needs assessment seems especially problematic. We have little to say on the topic of using tests for early childhood needs assessment save to echo Wargo et al's 1972 recommendation that standards and procedures need to be developed for use of tests in needs assessment*. For now, suffice it to note that whether one should use tests to assess students' needs or to assess programs' effects begs a much larger issue -- namely the purpose of evaluation. In the next chapter, we turn to consider this larger issue more directly.

---

*The need for developing such standards seems especially great in light of the tendency noted in section 5.4 for total tests to be rated as having higher technical quality than subtests. To identify students' particular needs, subtests keyed to specific skill areas should presumably be more important than total test scores that average students' performance over a broad range of skills.

equivalent to only the 9th percentile, or the 22nd NCE. In other words, as the Washington D. C. evaluators put it, "if a student falls asleep and learns nothing during first grade, he/she would be expected to lose 28 NCEs" (p.72). In contrast to this sharp change in 50th percentile performance during first grade, the Washington D. C. evaluators noted that the corresponding growth expectancies on CTBS Total Reading scores are 9.3 NCEs at grade 2; 8.7 at grade 3 and only 3.7 at grade 7. In short, the expected growth of children in terms of national norms (expressed either in the form of NCEs, percentiles, standard deviations, or stanines) is considerably larger for children having their first school experience than for children in later grades.

This phenomenon may help to explain the prevalence of exemplary programs at the early childhood level and the fact that in longitudinal evaluations of programs like Follow Through, more positive program effects are apparent at kindergarten and grade 1 levels than at later grade levels*. Since the statistical and educational significance of program effects are usually determined by comparing apparent gains to appropriate score distributions, the larger expected increase in children scores during early school experience may be the direct cause of more apparent program effects at this level.

The root causes of this phenomenon may be several - for example, having to do with the plasticity of early development, or the nature of early school experience. But whatever the causes of this phenomenon, it seems clear that it raises real problems for evaluating early childhood interventions in ways

---

* With regard to the early effects of Follow Through, see Haney (1977), pp 316-318. For elaboration on the idea that expected educational achievement growth in terms of norm distributions (or standardized growth expectancies) are greater for the early grades than later ones, see Stenner, et al (1978).

## VI. SUMMARY AND CONCLUSIONS

. In conclusion, let us draw back for a moment from the specific considerations treated up to now and ask a more fundamental question: Why evaluate? With respect to evaluations of Title I programs, a relevant answer is that, as noted in Chapter I, the United States Congress has in public law mandated the U. S. Commissioner of Education to provide for evaluations which "describe and measure the impact of programs and projects assisted under this Title." Beyond the legal mandate, however, one must consider the ends which evaluation of Title I ought to serve. As recounted in section 3.5, the point usually noted as distinguishing evaluation from basic research is that it is intended to inform the decision-making process.

If we take seriously this purpose of evaluation -- namely to aid in decision-making -- there are several ways in which we might look for direction to our evaluation efforts. We might, for example, review the history of Title I evaluations to find those studies which appear to have had most impact on past decision-making. On the assumption that future evaluations will likely serve decision-making best if they are similar to evaluations that have most influenced decision-making in the past, this strategy might make sense. This strategy would, however, seem to lead us to a rather perverse conclusion. In their summary of fifteen major Title I evaluations, Rossi et al, indicated that two seemed to have had unusual influence on decision-making. The Washington Research Project, they noted, "clearly affected the view of federal policymakers toward the usefulness of local education agency evaluations. (Rossi et al, 1977 p. 17). In the 1975 GAO report, "several recommendations were made to Congress, all of which were agreed to by USOE" (Rossi et al, p. 28). The large influence of these two reports seems odd because of all fifteen reports reviewed by Rossi, et al, these two receive the severest methodological criticisms. Rossi et al commented, for example, that "there was relatively little systematic data

collection" in the Washington Research Project study, and that the "USGAO report is seriously deficient in meeting fundamental scientific standards." (pp. 72, 148). So such a retrospective look for Title I evaluations exemplary in terms of informing decision-making would seem to suggest that future evaluations will more likely aid decision-making if they are methodologically "deficient" and not conducted by USOE! Such a conclusion not only seems perverse, but may also be somewhat wrong-headed. For it ignores the point noted in Chapter II, namely that if we are to improve evaluation's link to decision-making, we must consider the range of decisions within the power of the decision-makers who may apply what is learned from evaluations.

So a more sensible strategy to defining the practical purpose of Title I evaluation is to ask the following: Who are the decision-makers to be informed by Title I evaluations, and what sort of information do they want to inform their decisions? There are no clear-cut answers to these questions. In the Title I system, there are literally thousands of people whose decisions might be informed by Title I evaluations. Potential decision-makers exist at all three tiers of the Title I system; at the national, state and local levels. Relatively little attention seems to have been given to the types of evaluation information people at these levels want. Moreover, the few inquiries into information needs have not come to the same conclusions.

Gamel, et al (1975) reported on one effort to assess Title I information needs:

> The task of determining the amount and type of evaluation information needed at each level of the Title I administrative hierarchy was accomplished somewhat more rigorously at the national than at other levels. Interviews were held with educational policy makers in order to learn their information requirements regarding Title I operations and effects.
>
> Among those interviewed were a small number of federal legislators and senior professional staff members of all of the legislative committees and subcommittees involved with the ESEA Title I program.

> The interviews were unstructured to give the interviewers freedom to discuss any topic of concern and to prevent possible preconceptions of the research staff from shaping or limiting interview coverage. While this approach unquestionably had merit, it unfortunately precluded a precise tallying of attitudes regarding specific issues. It was evident, nonetheless, that for nearly all of those interviewed, the single most pressing concern was for some sort of measure of program impact on the children served. The comment was frequently encountered that after some ten years of program operation, it seemed only reasonable to have some objective and representative data regarding student outcomes.
>
> There was also almost universal concern for information about the number of children served, the adequacy with which their needs had been assessed, the services offered to them, and the costs associated with Title I projects and programs....
>
> No first-hand data were collected regarding the information needs of state, district, or local school personnel involved in the administration of Title I projects or programs. (pp. 5-6, emphasis added).

A somewhat different conclusion regarding information needs was reached by N.I.E. investigators. On the basis of their reading of Title I's legislative history, the N.I.E. researchers reported that there are three purposes to Title I which are "clearly discernible in the statute and accompanying formal Congressional statements":

(1) To provide financial assistance to school districts in relation to their numbers of low-income children and, within those districts, to the schools with the greatest numbers of low-income students

(2) To fund special services for low achieving children in the poorest schools.

(3) To contribute to the cognitive, emotional, social or physical development of participating students

(Frechtling, et al, 1977, p. x)

The N.I.E. researchers concluded:

> Early national evaluations of Title I considered only the third goal - contributing to children's development - and often rendered judgments on the efficacy of the program without accounting for the diverse ways in which LEAs had implemented it. Those evaluations overlooked some important truths about Title I: It has several objectives, and under it LEAs deliver a range of services with a variety of aims and emphases to a diverse set of beneficiaries. In contrast to earlier evaluations, therefore, NIE's strategy is designed to (1) provide clear information about what Title I is accomplishing toward achievement of each fundamental purpose, and (2) examine the implications of alternative ways of organizing the efforts of the Federal, state and local governments to achieve these purposes. (Frechtling, et al, 1977, p. xi)

A third investigation pertaining to Title I evaluation needs was recently carried out by the GAO:

> To obtain State and Local agencies' views on Federal education program evaluations and related matters, GAO sent questionnaires to education agencies in all States and the District of Columbia and to a statistical sample of local school districts throughout the Nation. (GAO, 1977, p.i).

The GAO questionaire was sent to state and local officials responsible for administering programs funded under Title I, III and VII of ESEA. The GAO reported on its findings as follows:

> Local and State evaluation reports on Federal elementary and secondary education program effectiveness are intended to provide information that local, State and Federal officials can use to make policy and program decisions. However, State and local officials see important differences in the types of evidence of program effectiveness that they and officials at other levels prefer.
>
> Better communication is needed among the three levels on information they need to facilitate policy and program decisions. Questionnaire results raise this question: Should all three levels be served by a reporting system based on the same reports?
>
> Although State officials view Office of Education program officials as being most impressed by standardized, norm-referenced test results, and local officials view State and Office of Education officials in the same manner, State and local officials say that they are not most impressed by such results.

> Local officials prefer broader, more diverse types of information on program results than just these test scores, and they are most impressed by improvements in curriculum and instructional methods and gains in the affective domain (likes, dislikes, interests, attitudes, motives, etc.) State officials are most impressed by results from criterion-referenced tests. (GAO, 1977, pp. iv-v)

We have quoted extensively from these three studies because the divergencies between their findings condition everything we will be able to conclude about the feasibility of developing a system to evaluate Title I early childhood programs. Specifically, all of our conclusions about the feasibility of developing an early childhood Title I evaluation system will be conditioned on the sorts of information needs a system is designed to meet. Moreover, as indicated above, available evidence suggests that information needs at the local, state and federal levels of the three-tiered Title I system may be so diverse as to be irreconcilable within a single evaluation system. This issue will be explored more fully in Phase II of Huron's investigation of the feasibility of developing models for evaluating early childhood programs funded under Title I. Also, we will address the issue of alternative users and uses of Title I evaluation information in our full feasibility report. For now, however, let us return to the narrower perspective adopted in foregoing chapters of this report, to summarize what's been learned from our review of the literature.

Trends in Early Childhood Education

Some of the trends identified in Chapter II have relevance for all levels of Title I early childhood programs. Three of those trends - toward individualization of programs, toward comprehensive services and ecological approaches to ECE, and toward coordination of ECE at the state level, all suggest that it may be increasingly difficult to identify specific educational programs which have been the object of most past educational evaluation

efforts. The increasing attention given to parents' involvement in their children's education, and its special relevance at the early childhood levels suggests that parents should be considered in evaluations of programs at this level. The trend towards mainstreaming seems to hold most relevance for first grade programs since Title I at the earlier levels -- especially at PK -- seems mainly to fund entire intervention programs wherein mainstreaming (or "pull-out"), is not a relevant issue.

Evaluation of Federal Early Childhood Education Programs

The five issues identified in Chapter III from our review of evaluations of federal programs in early childhood education seem to have several implications for Title I early childhood. The pattern of broad goals to intervention efforts seems to be more relevant to early childhood Title I efforts than later grade Title I interventions, but still ECTI efforts do not seem as broad as comprehensive intervention programs like Follow Through. The issue of long-range goals versus short-range evaluations of ECT-I will have to be faced. Pelavin and David (1977) have suggested that follow-up evaluations of compensatory education efforts often do not turn up the same results as simple fall-to-spring evaluations. Though their study covered only later grade Title I programs, their point is doubtless relevant to early childhood interventions -- as shown by the history of Head Start and Follow Through evaluations. Program definition raises several problems for evaluating early childhood Title I. Clearly, we need to know more about what kinds of ECTI programs are being funded. Also, however, we must ask whether the same sorts of evaluation procedures can possibly be appropriate for the very different kinds of programs encompassed under ECT I. The problems of uncertain findings and rare use of those findings both highlight the basic issue noted above, namely: What information needs and which decision-makers should an ECTI system aim to serve?

## Early Childhood Title I Programs

In our review in Chapter IV of what we know about ECTI programs, we found several things. Early childhood programs have received very little attention in past Title I evaluations - except in efforts to identify exemplary programs. Substantial resources seem to have been invested in Title I early childhood programs in at least 20 states around the country. Programs funded at Prekindergarten, Kindergarten and Grade 1, from what little we now know of them, appear to differ substantially. Selection and needs assessment of children to participate in ECT I programs raise special problems, both because of lack of appropriate needs assessment devices and because of unclear USOE regulations concerning appropriate selection procedures for ECT I programs. Parent involvement, though mandated for all Title I programs, seems to have special relevance to early childhood programs. The degree of support for parental involvement, seems however, to vary markedly in different states, and as was mentioned in Chapter 2, there is some disagreement in the literature over the appropriateness of emphasizing the participation or training aspects of parent involvement.

Previous evaluations of Title I programs revealed a host of problems. There was very little information that could be unambiguously aggregated across more than a handful of state Title I evaluation reports. States appeared to influence local Title I evaluation practices in widely differing degrees. Nevertheless, nearly 20 states evidenced intentions in their 1975-1976 reports to implement one of the three Title I evaluation models, although several commented on the difficulty of doing so for ECT I programs. Several states appear to have performed outcome evaluations at the first grade level, but relatively few appear to have done so at the PK and K levels.

Early Childhood Tests and Instruments

Many of the problems in conducting impact evaluations of early childhood programs have to do with tests and instruments for use with young children. Tests used for evaluation purposes were reviewed in Chapter V. Instruments used in FY 1976 evaluations of Title I often seem to have been parts of academic achievement test series frequently used at upper grade levels; namely the California Achievement Test, the Gates MacGinitie Reading Test, the Metropolitan Achievement Test, the Stanford Achievement Test, and the Wide Range Achievement Test. At the PK-K level, however, there seems to have been more tests of cognition and what are loosely called readiness tests. There are a variety of characteristics of young children that make traditional testing difficult. Nevertheless, there are a wide range of tests available for use with children in the PK-K range. Most of these appear, however, to have very poor technical quality. There is some indication that the tests most commonly used are those of types that have relatively better technical quality. Such tests, primarily those in the domains of cognition and academic readiness, appear to be the ones that typically have been used in demonstrating the effectiveness of exemplary early childhood programs. Use of such tests to show effectiveness of early childhood programs may have resulted in a relatively high proportion of exemplary early childhood programs in part because of test norms. Children's growth in terms of norm scores changes in the early years of schooling experiences is greater than in later grades.

## Conclusion

This paper marks the completion of only the first phase of Huron's investigation into the feasibility of developing models to evaluate early childhood education programs funded under Title I. Hence, it is too early to speak with much specificity about the form that such evaluation models might best take. Indeed, it is apparent that the crucial issue in designing any model or system for evaluating ECT I programs is the information needs which are to be served. Available evidence indicates that information needs expressed by people involved in Title I at the local, state and national levels are not the same. This suggests that no one model or system can serve these different information needs equally well. As mentioned, these issues will be treated more fully in future phases of our work.

Nevertheless, in conclusion, let us draw attention to three findings of our literature review which seem to hold special relevance to our continuing work. Our review shows that there is a real paucity of descriptive information available on the sorts of early childhood programs funded under Title I. The second phase of our project, which entails field visits to such programs in ten states around the country, will help remedy this problem for the purposes of the present project. Still, our review of past evaluations of early childhood education programs shows clearly that there is a larger problem involved. Specifically, there is a real need for on-going evaluations of ECT I to include more descriptive information on exactly what kinds of early childhood programs are being funded under Title I.

Second, from the descriptive information available, it is apparent that a wide variety of early childhood programs are being funded under Title I. At the pre-kindergarten level, programs seem usually to comprise relatively broad intervention efforts or what might be called whole-program efforts. At the first-grade level, in contrast, Title I funds seem more often to be used to pay for supplementary services which are provided above and beyond childrens'regular school programs. At the kindergarten level, both whole-program and supplementary-program interventions were apparent. These differences in the structure of interventions funded at these three levels, together with the differences in the goals of programs at these levels, suggests that no single evaluation scheme will serve these levels equally well.

Third, our review of tests and instruments for use at the early childhood level, showed that there is a real paucity of high quality instruments for use with young children. Moreover, our review suggests that judgments regarding the effectiveness of early childhood programs may be confounded with the quality of the instruments by which their effectiveness is judged. This suggests that there is a real need for evaluation models or at least guidelines by which early childhood programs may be assessed in areas where no adequate instruments are available.

## APPENDIX I: STATE TITLE I EVALUATION REPORTS FOR 1975-1976 BY STATE FOR 1975-1976*


ALABAMA:

Nettles, D. ESEA, Title I, Public Law 89-10 (as amended) Annual Evaluation Report, 1975-1976. Montgomery, Alabama: Alabama Department of Education, State Office Building, 1975.

Nettles, D., and Anderson, R. Diagnostic and prescriptive teaching. A Self-Directed Program for Staff Development. ESEA Title I Public Law 89-10. Alabama State Department of Education, Montgomery, Alabama: 1976.

ALASKA:

State Department of Education. ESEA Title I 1976 Evaluation. Alaska: 1976.

ARIZONA:

Arizona State Department of Education. ESEA Title I in Arizona, Fiscal Year 1976. Arizona: State Department of Education, 1976.

Arizona State Department of Education. Parents: Creative Sources of Learning. Tucson: Tucson Public Schools, 1976.

Slaughter, H. Parent and Child Education Project, ESEA Title I Annual Evaluation Report, 1975-1976. Tucson: Tucson Public Schools, Research and Evaluation Department, 1976.

ARKANSAS:

Andrew, Dean, and Program Staff. 1976 Evaluation Report, Title I, ESEA. Arkansas: Department of Education, Title I, ESEA Program Staff, 1976.

BUREAU OF INDIAN AFFAIRS:

Bureau of Indian Affairs. Bureau of Indian Affairs, Title I, ESEA Annual Evaluation Report, Fiscal Year 1976. Albuquerque, New Mexico: Office of Indian Education Programs, Indian Education Resources Center, 1976.

CALIFORNIA:

California State Department of Education. Evaluation Report of ECE, ESEA Title I, and EDY, 1975-1976. Sacremento, California: California State Department of Education, 1977.


---

* In this appendix, we have listed state Title I Evaluation Reports for 1975-1976 by state. For the few states from which the 1975-1976 reports were not yet available, previous years' reports are listed.

COLORADO:

Cheuvront, R. The Elementary and Secondary Education Act, Title I (Public Law 98-10) Colorado Annual Evaluation Report. Denver, Colorado: Colorado State Department of Education, 1976.

CONNECTICUT:

Connecticut State Department of Education. Compensatory Education in Connecticut: Fiscal Year 1976. Hartford: 1977.

Connecticut State Department of Education. Connecticut Compensatory Education Programs Annual Evaluation Report 1975-1976 Programs for Neglected and Delinquent Children (Report Summary). Hartford: 1977.

Connecticut State Department of Education. Reading and Mathematics Skills Achievement in Title I and S.A.D.C. Compensatory Programs, 1975-1976 (Report Summary). Hartford: 1977.

Connecticut State Department of Education. Summary of Achievement Analysis (for Parent Advisory Councils). Hartford: November 30, 1977.

Bordonaro, A. Achievement Analysis, 1975-1976 - Compensatory Reading and Mathematics Programs Conducted Under Title I of the Elementary and Secondary Education Act and the State Act for Disadvantaged Children, Fiscal Year 1976. Conn: Connecticut State Department of Education, 1977.

Bordonaro, A. Connecticut Annual Evaluation Report, 1975-1976 Compensatory Education Programs. Connecticut State Department of Education, 1977.

Whitelaw, D. Connecticut Compensatory Education Programs-Programs for Neglected and Delinquent Children Supported by Title I of the Elementary and Secondary Education Act, Annual Evaluation Report 1975-1976. Connecticut State Department of Education, Division of Instructional Services, Bureau of Elementary and Secondary Education, November, 1977.

DISTRICT OF COLUMBIA:

IBEX, Incorporated and Roy Littlejohn Associates, Inc. Evaluation of the ESEA Title I Program of the Public Schools of the District of Columbia 1975-1976. Final Evaluation Report. Washington D.C.: IBEX Incorporated and Roy Littlejohn Associates, Inc.,1976.

DELAWARE:

Corkel, W. Annual Report of ESEA Title I P.L. 89-10 FY 1976. Dover, Delaware: Delaware Department of Public Instruction, 1976.

FLORIDA:

Office of Compensatory Education. Florida Title I Evaluation Report Fiscal Year 1976: Elementary and Secondary Education Act of 1965. Tallahassee, FL: Florida Department of Education, January 1977. Series 77-10.

GEORGIA:

Moore, S., and White, L. Title I in Georgia Annual Evaluation Report 1974-1975. State of Georgia for Projects Supported by Title I, ESEA Funds. Atlanta: Division of Planning and Evaluation. Office of Administrative Services. Georgia Department of Education, 1977.

HAWAII:

Office of Instructional Services/Compensatory Education Section. State Summary Evaluation Report ESEA Title I, P.L. 89-10, School Year 1975-76. Hawaii Department of Eduaction, March 1977.

IDAHO:

Idaho State Department of Education. Title I ESEA Evaluation Report Fiscal Year 1976. Boise, Idaho; 1976

ILLINOIS:

West, R. FY 76 Evaluation Report on Title I 89-10. Illinois: State Board of Education, Illinois Office of Education, November 1976.

INDIANA:

Division of Compensatory Education. Annual Evaluation Report Title I ESEA Fiscal Year 1976. Indianapolis, Indiana: State of Indiana Department of Public Instruction, 1976.

IOWA:

Department of Public Instruction. Exemplary Project Proposal Cedar Rapids Community Schools.

Department of Public Instruction. Iowa Annual Evaluation Report for Fiscal Year 1976. Title I of Public Law 93-380 The Elementary and Secondary Education Act. Des Moines, Iowa: State of Iowa Department of Public Education, 1976.

KENTUCKY:

Office of Special Programs and Services. Evaluation Report 1975-1976 Kentucky ESEA Title I Program. Kentucky: Office of Special Programs and Services with the Divisions of Compensatory Education and of Computer Services, 1976.

KANSAS:

Central Research Corporation. (E.S.E.A.) Title I in Kansas. Annual Evaluation Report 1975-1976.

LOUISIANA:

Henslee, B.,Gerace, V.,and Norris, J. State Annual Evaluation Report for Fiscal Year 1976, Public Law 89-10, Title I, ESEA. Louisiana State Department of Education, 1976.

MAINE:

Title I Elementary and Secondary Education Act. Maine Eleventh Annual Report, FY 1976.

MARYLAND:

Maryland State Department of Education. Annual Evaluation Report on Programs Funded Under ESEA Title I in Maryland for Fiscal Year 1976. Maryland: State Department of Education, Division of Compensatory, Urban, and Supplementary Programs, 1977.

MASSACHUSETTS:

State Department of Education. Annual Evaluation Report: Title I, ESEA, FY 1975. Boston: Department of Education, Bureau of Curriculum Services, 1975.

State Department of Education. Annual Report, ESEA and Title I, 1974. Boston: Department of Eduaction, Bureau of Curriculum Services, 1975.

MICHIGAN:

Department of Education. A Description and Evaluation of 1974-1975 Regular School Year ESEA. Title I Programs in Michigan. Michigan: Department of Education, Research, Evaluation, and Assessment Services, 1975.

MINNESOTA:

State Board of Education. Minnesota ESEA Title I evaluation Report Fiscal Year 1976. Minnesota: State Board of Education, Department of Education, 1977.

MISSISSIPPI:

State Department of Education. Annual Evaluation Report. Title I, ESEA Regular School Year 1975-1976. Mississippi: State Department of Education, Title I, ESEA Administrative Unit, 1977.

MISSOURI:

State Department of Elementary and Secondary Education. State of Missouri Annual Evaluation Report for Title I, ESEA. Jefferson City, Missouri: Missouri State Department of Elementary and Secondary Education, 1976.

MONTANA:

Office of Public Instruction. State Annual Evaluation Report for 1975-1976 Title I Elementary and Secondary Education Act of 1965 Public Law 89-10, as amended. Helena, Montana: State of Montana Office of Public Instruction, 1977.

NEVADA:

Department of Education. ESEA Title I Annual Evaluation Report State of Nevada 1975-1976. Nevada: Department of Education, 1977.

NEBRASKA:

State Department of Education. Annual Evaluation Report-1976. Title I ESEA. Nebraska: State Department of Education, 1977.

NEW HAMPSHIRE:

Department of Education. Title I ESEA Fiscal Year, 1976. Evaluation Summary Report. New Hampshire: Department of Education, 1977.

NEW JERSEY:

State Department of Education. ESEA Title I: Annual Evaluation Report, Fiscal Year 1975. New Jersey: State Department of Education, 1976.

NEW MEXICO:

Albuquerque Public Schools. A Non-testing Method for Identifying Title I Participating Children. Albuquerque, New Mexico: Albuquerque Public Schools,Title I Planning and Evaluation, 1977.

Albuquerque Public Schools. Title I-Past and Present. Albuquerque Public Schools, Title I Planning and Evaluation, 1977.

Wengs, K. Title I in New Mexico. Annual Evaluation Report for Fiscal Year Ending June 30, 1976. Santa Fe, NM: New Mexico State Department of Education, 1976.

NEW YORK:

University of the State of New York. A Survey of ESEA, Title I Participants' Successful Performance in Criterion-Referenced Objectives: Reading and Mathematics. Albany: University of the State of New York, the State Education Department, Bureau of School and Categorical Programs Evaluation, 1977.

University of the State of New York. ESEA Title I Evaluation Report, New York State 1975-1976. Albany: University of the State of New York, the State Education Department, Bureau of School and Categorical Programs Evaluation, 1977.

University of the State of New York. Preliminary Report of Findings-Evaluation of the New York State Experimental Prekindergarten Program. Albany, New York: The University of the State of New York, State Edication Department, Division of Research, Prekindergarten Evaluation Unit, 1977.

NORTH CAROLINA:

Division of Compensatory Education. Elementary and Secondary Education Act of 1965- Title I. North Carolina Annual Evaluation Report Fiscal Year 1976. Raleigh, NC: North Carolina Department of Public Instruction, State Department of Public Instruction, 1976.

NORTH DAKOTA:

Department of Public Instruction. North Dakota Title I ESEA Evaluation Report Fiscal Year 1976. Bismark, ND: Department of Public Instruction, Title I, ESEA Program Staff, 1977.

OHIO:

State of Ohio Department of Education. The Eleventh Annual Evaluation Report of Title I in Ohio Fiscal Year 1976. Columbus, OH: State of Ohio Department of Education, 1976.

OKLAHOMA:

State Department of Education. Annual Evaluation Report Title I 1975-1976. State of Oklahoma Department of Education, Compensatory Education, Title I ESEA Section, 1977.

OREGON:

Hunt, B. Oregon Title I Annual Evaluation Report FY 1976. Salem, OR: Oregon Department of Education, 1977.

PENNSYLVANIA:

Department of Education. Part I Annual Evaluation of Title I-ESEA Programs. Harrisburg, PA: Commonwealth of Pennsylvania Department of Education, Division of Compensatory Programs, Bureau of Special and Compensatory Education, 1976.

Department of Education. Part II. Annual Evaluation Analysis of ESEA Title I Test Data Submitted by Pennsylvania School Districts. Harrisburg, PA: Commonwealth of Pennsylvania Department of Education, Division of Special and Compensatory Programs, Bureau of Special and Compensatory Education, 1976.

PUERTO RICO:

Department of Education. State Annual Evaluation Report, Title I, ESEA Fiscal Year, 1976. Hato Rey, PR: Commonwealth of Puerto Rico Department of Education, Title I Evaluation.

RHODE ISLAND:

DeVito, P. Compensatory Education (Title I, ESEA and State Section 4) Programs in Rhode Island. Eleventh Annual Evalaution Fiscal Year 1975-1976. Providence, RI: Rhode Island Department of Education, Bureau of Grants and Regulations, 1976.

SOUTH CAROLINA:

Department of Education. ESEA Title I Annual Evaluation Report FY 1976. Columbia, SC: South Carolina Department of Education, Office of Federal Programs, 1976.

SOUTH DAKOTA:

Department of Education and Cultural Affairs. Evaluation 1976 Title I. Pierre, SD: Department of Education and Cultural Affairs, Division of Elementary and Secondary Education, 1977.

TENNESSEE:

State Department of Education. 1976 ESEA Title I Annual Evaluation Report. Tennessee: State Department of Education, 1976.

TEXAS:

Texas Education Agency. Annual Report of Programs for the Disadvantaged in Texas 1975-1976. Austin, TX: Texas Education Agency, Division of Assessment and Education, 1976.

UTAH:

State Board of Education. Utah Annual Evaluation Report Title I, ESEA. Salt Lake City, UT: State Board of Education, Division of Technical Assistance, 1976.

VERMONT:

Knight, M. Jr. Elementary and Secondary Education Act Title I- P.L. 89-10 in the State of Vermont. Annual Report 1975-1976. Vermont: Department of Education, Division of Federal Programs, Compensatory Education, 1976.

VIRGIN ISLANDS:

Department of Education, ESEA Title I Teacher Trainee Program Evaluation Report 1975-1976 School Year. St. Thomas, VI: Department of Education, Division of Planning, Research, and Evaluation, 1976.

VIRGINIA:

State Department of Education. Annual Evaluation Report, Title I Elementary and Secondary Education Act of 1965, Fiscal Year 1976. Richmond, Virginia: Department of Education, 1976.

WASHINGTON:

Cupp, M., and Rasp, A. Jr. Annual Report and Evaluation. ESEA Title I (P.L. 89-10) Fiscal Year 1976. Olympia, WA: Department of Public Instruction, 1977.

WEST VIRGINIA:

Board of Education. State Annual Evaluation Report-Fiscal Year 1975, Title I. West Virginia: Board of Education, 1975.

WISCONSIN:

Department of Public Instruction. State of Wisconsin Annual Evaluation Report: Title I, Elementary and Secondary Education Act of 1965, Fiscal Year 1976. Wisconsin: Department of Public Instruction, Title I Unit, Division of Instructional Services, 1976.

Department of Public Instruction. Title I, ESEA Directory of Preschool Projects, 1975. Wisconsin: Departmenr of Public Instruction, 1976.

Department of Public Instruction. Wisconsin Annual Evaluation, Title I - ESEA, Fiscal Year 1974. Wisconsin: Department of Public Instruction, Title I Unit, Division of Instructional Services, 1975.

WYOMING:

State Department of Education. Wyoming Annual Evaluation Report Title I, ESEA Fiscal Year 1976. Cheyenne, Wyoming: State Department of Education, Instructional Services Division, 1976.

## APPENDIX II: GENERAL BIBLIOGRAPHY

Abbott, M.; and Crane, J. Journal of School Psychology, 15, 1977, 118-128.

Abraham, W. The Early Years: Prologue to Tomorrow. Exceptional Children, 42, 1976, 330-335.

Anastasi, A. Personal Communication, April 12, 1978.

Anastasi, A. Psychological Testing. (4th Ed.) New York: MacMillan, 1976.

Anderson, R., et al. "Pardon Us, But What Was That Question Again?: A Response to the Critique of the Follow Through Evaluation." Harvard Educational Review, 48:2, May 1978, 161-170.

Baker, E. Evaluation of the California Early Childhood Education Program. Executive Summary. California: UCLA Graduate School of Education, 1977.

Baker, E. Evaluation of the California Early Childhood Education Program, Vol. 1, The Evaluation of the Early Childhood Education Program. California: The UCLA Graduate School of Education, Center for the Study of Evaluation, December, 1976.

Baker, E.; Keesling, J.; and Burstein, L. Evaluation of the California Early Childhood Education Program, Volume II, An Audit Report on the Activities of the State Department of Education Related to the Early Childhood Education Program. California: UCLA Graduate School of Education, Center for the Study of Evaluation, December, 1976.

Barker, G.; and Hurd, L. Preprimary Enrollment, October, 1970. Washington, D. C.: U. S. Government Printing Office, 1971.

Barnes, D. Recent Trends in Education. Teacher Education 11, 1976, 10-22.

Bauch, J.; and Shearron, G. "The Elementary School of the Late 70's". ERIC System: ED 040 922, 1960.

Beveridge, L. Group Interaction Analysis: What? Who? Why? The George Washington University. December 4, 1972, unpublished manuscript.

Biber, B.; Shapiro, E.; and Wickens, D. Promoting Cognitive Growth: A Developmental-Interaction Point of View. Washington, D. C.: National Association for the Education of Young Children, 1971.

Bijou, S. Child Development: The Basic Stage of Early Childhood. Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1976.

Bloom, B. Stability and Change in Human Characteristics. New York: John Wiley and Sons, 1964.

Bowman, G. 1972 Analysis of Communication in Education: Follow Through Program. The Bank Street College of Education, New York: 1972.

Boyer, E.; Simon, A., and Karafin, G. Measures of Maturation: An Anthology of Early Childhood Observation Instruments. Philadelphia, Pennsylvania: Research for Better Schools, 1973.

Branch, C.; and Overly, N. Illustrative Descriptions of Two Early Childhood Education Programs. Educational Leadership, 28, 1971, 821-826.

Braun, S.; and Edwards, E. History and Theory of Early Childhood Education. Worthington, Ohio: C. A. Jones, 1972.

Bricker, D.; and Bricker, W. A Developmentally Integrated Approach to Early Intervention. Education and Training of the Mentally Retarded, 12, 1977, 100-108.

Bronfenbrenner, U. A Report on Longitudinal Evaluations of Preschool Programs. Volume II: Is Early Intervention Effective? DHEW Publication No. (OHD) 74-25, 1974.

Bureau of Child Development and Parent Education. Children's Spaces: A Design for the 70's Planning Educational Facilities in the Elementary School for Very Young Children. Albany, New York: The University of the State of New York Press, no date.

Bureau of Child Development and Parent Education. Guides for Selection of Indoor and Outdoor Equipment and Materials, Suggestions for Administrators and Teachers. Albany, New York: The University of the State of New York, The State Education Department, 1967.

Bureau of Child Development and Parent Education. Pre-K Day Care is School Plus. Albany, New York: The University of the State of New York Press, no date.

Buros, O. Seventh Mental Measurement Yearbook. New Jersey: Gryphon Press, 1972.

Buros, O. Personality Tests and Reviews. New Jersey: Gryphon Press, 1970.

Buros, O. Reading Tests and Reviews. New Jersey: Gryphon Press, 1968.

California State Department of Education. Report of the Task Force on Early Childhood Education. Sacremento, California: Bureau of Publication, 1972.

Cannara, A., Chiang, A., Bessey, B. The Title I Evaluation and Reporting System Computer Program: Documentation and Use. RMC Research Corporation, Mountain View, California, June 10, 1977.

Carpenter, P. (ed.) The Education Index, IV. 1938-1941, N.Y.: The H. W. Wilson, Co., 1941.

Carpenter, P. (ed.) The Education Index, VII, 1947-1950, N.Y.: The H. W. Wilson Co., 1950

Center for Educational Research and Innovation. Developments in Early Childhood Education. Paris: Organization for Economic Cooperation and Development, 1975.

Chaffee, J., Jr. (ed.) Compact 7. July/August, 1973. Denver Colorado: Education Commission of the States, 1973.

Chum, K., Cobb, S., and French, J. Measures for Psychological Assessment: A Guide to 3000 Original Sources and Their Applications. Ann Arbor, MI: Survey Research Center, 1975.

Cicirelli, and others. The Impact of Head Start: An Evaluation of the Effects of Head Start on Children's Cognitive and Affective Development. Report to the Office of Economic Opportunity by Westinghouse Learning Corporation and Ohio University. Washington, D. C.: Office of Economic Opportunity, 1969.

Cicirelli, V., Evans, J. and Schiller J. "The Impact of Head Start: A Reply to the Report Analysis. Harvard Educational Review, 40:1, February 1970.

Coleman, J.; and Sandoval, C. Pre-primary Education: Needs, Alternatives and Costs, 1971-1980. Washington, D. C.: Education and Public Affairs, 1971.

Collis, K. F. A Study of Concrete and Formal Operations in School Mathematics: A Piagetian Viewpoint. Victoria, Australia: Australian Council for Educational Research, 1975.

Council of Europe. Problems in the Evaluation of Preschool Education. Strasbourg: Documentation Center for Education in Europe, 1975.

Cronbach, L. Test Validation. In R. L. Thorndike (Ed.), Educational Measurement, (2nd Edition), Washington, D. C.: American Council on Education, 1971.

Datta, L. The Impact of the Westinghouse - Ohio Evaluation on the Development of Project Head Start: An Examination of the Immediate and Longer-Term Effects and How They Came About. In C. A. Abt. (Ed.), The Evaluation of Social Programs, Sage Publications, 1976.

Datta, L. Parent Involvement in the United States, in Center for Educational Research and Innovation (ed.), Developments in Early Childhood Education. Paris: Organization for Economic Cooperation and Development, 1975, 38-48.

Department of Health, Office of Education. Special Educational Needs of Migratory Children: Grants to State Educational Agencies. Federal Register, Volume 42, No. 154, Part II, Wednesday, July 13, 1977.

Department of Health, Education and Welfare, Office of Education. Grants to State Agencies for Programs to Meet the Special Educational Needs of Children in Institutions for Neglected or Delinquent Children: Interim Regulations. Federal Register, Volume 42, No. 70, Part II, Tuesday, April 12, 1977.

Department of Health, Education and Welfare, Office of Education. Financial Assistance to Local Educational Agencies to Meet the Special Educational Needs of Educationally Deprived, Neglected and Delinquent Children. Federal Register, Vol. 41, No. 189, Part IV, Tuesday, September 28, 1976.

Divoky, D. Mainstreaming: Moving special education out of the basement. Compact, 10, 1976, 2-5.

Dunlop, K. Mainstreaming: Valuing diversity in children. Young Children, 32, 1977, 26-32.

Ebel, R. Must All Tests Be Valid. American Psychologist, 16, (1961), pp. 640-47.

ERIC Clearinghouse on Early Childhood Education. Public Policy and the Health, Education and Welfare of Children: An Abstract Bibliography. ERIC/ECE Catalog, No. 160, April, 1977.

Education Commission of the States. Early Childhood Programs in the States: Report of a December 1972 Conference. Denver, Colorado: Education Commission of the States, 1972

Education Commission of the States. The Persistence of Preschool Effects. A National Collaborative Study by Twelve Research Groups. New York: Community Service Laboratory, Cornell University, September, 1977.

Ehrenreich, J. (ed.) The Education Index. XIX, 1968-1969, NY: H. W. Wilson Co. 19

Ehrenreich, J. (ed.) The Education Index. XXII, 1972-1973, NY: H.W. Wilson Co., 19

Elkind, D. Early Childhood Education in the seventies: What can we learn from past mistakes? Contemporary Education, 45, 1974, 254-260.

Elkind, D., and Lyke, N. Early education and kindergarten: competition or cooperation? Young Children , 30, 1975, 393-401.

Emrick, J., Sorenson, P., and Stearns, M. Interim Evaluation of the National Follow Through Program 1969-1971 (Technical Report to USOE). California: Stanford Research Institute, 1973.

Epps, E. Situational Effects in Testing. In Miller, L. The Testing of Black Students, Englewood Cliffs, NJ: Prentice-Hall, 1974.

Evans, E. Contemporary Influences in Early Childhood Education. N.Y.: Holt, Rinehart, and Winston, Inc., 1975.

Fleiss, B. The continuity of education, from prekindergarten to primary. Young Children, 22, 1966, 78-83.

Fortune, J., and Hutchinson, T., et al. Data Analysis of the 1969-1970 Survey of Compensatory Education (Title I), Volume I. Amherst, MA: University of Massachusetts, n.d.

Frazier, A. (ed.) "The New Elementary School". ERIC System: ED 017 341, 1968.

Frechtling, J. The Effects of Services on Student Development. Washington, D.C.: National Institute of Education, September 30, 1977.

Freud, A. Normality and Pathology in Childhood. New York: International Universities Press, 1965.

Gagne, R. Observations of school learning. Educational Psychologist, 10, 1973, 112-116.

Gallagher, J. Current trends in special education in the United States. International Review of Education, 20, 1974, 277-297.

Gamel, N., Tallmadge, G., Wood, C., and Binkley, J. State ESEA Title I Reports: Review and Analysis of Past Reports, and Development of a Model Reporting System and Format. California: RMC Corporation, October, 1975.

GAO. Report to the Congress by the Comptroller General: Assessment of Reading Activities Funded Under the Federal Program of Aide for Educationally Deprived Children. Washington, D. C.: U. S. General Accounting Office, December, 1975.

Garfinkel, J. Studies in Ethnomethodology. Englewood Cliffs, New Jersey: Prentice Hall, 1967.

General Accounting Office. Problems and Improvements in Evaluating Office of Education Programs. Washington, D.C.: General Accounting Office, 1977.

General Electric Company. TEMPO: Evaluation of Title I ESEA Compensatory Education. Santa Barbara, California: General Electric Co. - TEMPO, 1971.

General Electric Company. A Survey and Preliminary Analysis in Elementary-Secondary Education: Final Report. General Electric, 1967.

Gilstrap, R. Early childhood education in the 70's: A view from NEA center. Peabody Journal of Education, 48, 1971, 153-157.

Glass, G. V., et al. Data Analysis of the 1968-1969 Survey of Compensatory Education (Title I). Washington, D. C.: USOE, August, 1970. (ED 057 146).

Glass, G., and Smith, J. "Pull Out" in Compensatory Education. A Paper presented to the Office of the Commissioner, U. S. Office of Education, 1977.

Goldman, B., and Saunders, S. Directory of Unpublished Experimental Measures, Volume I. New York: Behavioral Publications, 1974.

Golladay, M. The Condition of Education, V. 3, Part 1. Washington, D. C.: U. S. Government Printing Office, 1977.

Goodlad, J., and Shane, H. (eds.) The Elementary School in the United States: The Seventy-Second Yearbook of the National Society for the Study of Education. Chicago, Illinois: University of Chicago Press, 1973.

Gordon, E. W., and Konrelokos, J. Utilizing Available Information from Compensatory Education and Surveys. June, 1971. (ED 055 664).

Gordon, I. (ed.) Early Childhood Education: The Seventy-first Yearbook for the National Society of the Study of Education, Part II. Chicago, Illinois: The University of Chicago Press, 1972.

Gordon, I. Parental Involvement in Compensatory Education. Urbana, Illinois: University of Illinois Press, 1971.

Grant, W., and Lind, C. Digest of Education Statistics, 1975. Washington, D. C.: U. S. Government Printing Office, 1976.

Grotberg, E. Review of Research: 1965-1969. Washington, D. C.: Project Head Start, United States Office of Economic Opportunity, 1969.

Guthrie, P., and Horne, E. Measures of Infant Development: An Annotated Bibliography. New Jersey: Educational Testing Service, Head Start Collection, 1971.

Guthrie, P., and Horne, E. Measures of Social Skills: An Annotated Bibliography. New Jersey: Educational Testing Service, Head Start Collection, 1971.

Guthrie, P., and Horne, E. School Readiness Measures: An Annotated Bibliography, New Jersey: Educational Testing Service, Head Start Collection, 1971

Haney, W. The Follow Through Planned Variation Experiment. Volume V: A Technical History of the National Follow Through Evaluation. Cambridge, Mass: Huron Institute, 1977.

Haney, W. "The Follow Through Planned Variation Experiment: Statistical Conclusion Validity?" Paper presented at the Annual Meeting of the Joint Dissemination Review Panel: Toronto, Canada, March 28, 1978.

Hawkridge, D., Chalupsky, A., and Roberts, A. A Study of Exemplary Programs for the Education of Disadvantaged Children. Palo Alto: American Institutes for Research, 1968 (ERIC System: ED 023 776, ED 023 777).

Harty, H., and Roberts, A. Recent trends to keep up with. Peabody Journal of Education, 51, 1973, 64-67.

Hendrick, J. The Whole Child: New Trends in Early Education. St. Louis, Mo: The C. U. Mosby Co., 1975.

Henny, N. (ed.) The Forty-Sixth Yearbook of the National Society for the Study of Education: Part II, Early Childhood Education. Chicago, Ill: University of Chicago Press, 1947.

Hertz, T. The Impact of Federal Early Childhood Programs on Children, Executive Summary. July, 1977.

Hertz, T. Mangus, S., Mann, A., and Wagner, C. Toward Interagency Coordination: FY 1976 Federal Research and Development on Early Childhood Sixth Annual Report. Washington, D. C.: The Social Research Group, George Washington University, 1977.

Hess, R., Block, M. Costello, J., Knowles, R., and Largay. Parent involvement in early education. In E. Grotberg (ed.) Day Care: Parent Involvement Resources for Decision. Washington, D. C.: United States Government Printing Office, 1971, 265-298.

Hewitt, H. (ed.) The Education Index. 1975-1976, N. Y.: The H. W. Wilson Co, 1976.

Heyneman, S. P. The Frequency of Tests and Measures Utilized in Proposals for Federal Government Research and Development on Children and Adolescents in FY 1974. Washington, D. C.: Social Research Group, George Washington University, 1975.

Heyneman, S., and Mintz, F. The Frequency and Quality of Measures Utilized in Federally Sponsored Research on Children and Adolescents. Washington, D. C.: Social Research Group, George Washington University, 1976.

Hick, T., Kukuk, S. Horan, M. and Irvine, D. Effects of Prekindergarten on Three Measures of Cognitive Development. Albany, New York: The University of the State of New York, State Department of Education, Division of Research, Technical Paper No. 3, May 25, 1977.

Higgins, F. "Mainstreaming: An overview and update." ERIC System, ED 122 466, 1976.

Hill, P. Administration of Compensatory Education: A Report from the National Institute of Education. Washington, D. C. United States Department of Health, Education and Welfare, 1977.

Hill, P. Compensatory Education Services. A Report from the National Institute of Education. United States Department of Health, Education and Welfare, 1977

Hill, P. Demonstration Studies of Funds Allocation Within Districts. A Report from the National Institute of Education. Washignton, D. C.: United States Department of Health, Education and Welfare, 1977.

Hill, P. The Effects of Services on Student Development. A Report from the National Institute of Education. Washington, D. C.: United States Department of Health, Education, and Welfare, 1977.

Hill, P., and Rothberg, I. Evaluating Compensatory Education: An Interim Report on the National Institute of Education Compensatory Education Study. Washington, D. C.: United States Department of Health, Education and Welfare, 1976.

Hill, P., and Rothberg, I. Title I Funds Allocation: The Current Formula. Washington, D. C.: The National Institute of Education, September, 30, 1977.

Himelrick, J,, and Aitken, S. An Introduction to State Capacity Building: Planning, Management, and Delivery of Child and Family Services. Denver, Colorado: National Association of State Directors of Child Development and the Education Commission of the States, Report No. 98, 1976.

Hobbs, N. The Futures of Children: Categories, Labels and Their Consequences. Washington, D. C.: Jossey-Bass, 1975.

Hodges, W. "The Work of the Follow Through Experience." Harvard Educational Review, 48:2, May 1978, 186-192.

Hoepfner, T. (ed.) Measuring Normal Affective States in Children. Los Angeles, California: 1970.

Hoepfner, R., and others. CSE Elementary School Test Evaluations. Los Angeles, California: Center for the Study of Education and Early Childhood Research Center, UCLA, 1976.

Hoepfner, R., Stern, C., and Nummedal, S. (eds.) CSE-ECRC Preschool/Kindergarten Test Evaluations. Los Angeles, California: Center for the Study of Evaluation and Early Childhood Research Center, UCLA, 1971.

House, E., Glass, G., McLean, L., and Walker, D. No Simple Answer: Critique of the "Follow Through Evaluation". 1977.

Hunt, J. Reflections on a decade of early education. Journal of Abnormal Child Psychology, 3, 1975, 275-330.

Hunt, J. Intelligence and Experience. New York: Ronald, 1961.

Hymes, J. Emerging patterns in early childhood education. Young Children, 22, 1967, 158-164.

Illingsworth, R. S. The Development of the Infant and the Young Child, Normal & Abnormal. London: Churchill & Livinstone, 1972.

Jaeger, R. "On the Legitimacy of Combining Achievement Test Data Through NCE Scaled Scores." North Carolina: University of North Carolina at Greensboro, and University of South Florida, 1977, unpublished ms.

Jencks, C. Inequality: A Reassessment of the Effect of Family and Schooling in America. New York: Basic Books, 1972.

Johnson, O. Tests and Measurements in Child Development: Handbook II, Volumes 1 and 2. San Francisco, California: Jossey-Bass, 1976.

Johnson, O., and Bommarito, J. Tests and Measurements in Child Development: A Handbook. San Francisco: California: Jossey-Bass, 1976.

Karnes, M., with Hodgins, A., Teska, J., and Keik, S. Research and Development Program on Preschool Disadvantaged Children: Investigations of Classroom and At-Home Interventions. Washington, D. C.: Department of Health, Education and Welfare, 1969.

Kaskowitz, D., and Norwood, C. A Study of the Norm-Referenced Procedure for Evaluating Project Effectiveness as Applied in the Evaluation of Project Information Packages. California: Stanford Research Institute, 1977.

Kilmer, S., and Weinberg, R. The nature of young children and the state of early education: reflections from the Minnesota Round Table. Young Children, 30, 1974, 60-67.

King, I. Preprimary Enrollment, October, 1974. Washington, D. C. U. S. Government Printing Office, 1975.

Lazar, I., Hubbell, V., Murray, H. Rosche, M., and Royce, J. The Persistence of Preschool Effects. (Report for the Education Commission of the States), Ithaca, New York; Community Service Laboratory, N. Y. S. College of Human Ecology, Cornell University, 1977.

Lee, D. Views of the Child. In J. Goodlad and H. Shane (eds.) The Elementary School in the United States: The Seventy-Second Yearbook of the National Society for the Study of Education. Chicago, Illinois: University of Chicago Press, 1973, 138-173.

Leeper, S. Reflections/Directions. Childhood Education, 53, 1976, 4-8.

Lidz, C. Issues in the psychological assessment of preschool children. Journal of School Psychology, 15, 1977, 129-135.

Liederman, D. Why the states must lead. Compact, 7, 1973, 4-5

Lindsey, W. Instrumentation of OCD Research Projects on the Family. Washington, D. C.: The Social Research Group, The George Washington University, 1975.

Loret, P., Seder, A., Bianchini, J., and Vale, C. Equivalence and Norms Tables for Selected Reading Achievement Tests. New Jersey: Educational Testing Service, 1974.

Mann, A., Harrell, A., Hurt, M. A Review of Head Start Research Since 1969 and An Annotated Bibliography. Washington, D. C.: The Social Research Group, The George Washington University, 1977.

Mardell, C., and Goldenberg, D. "Instruments for Screening Prekindergarten Children." Springfield, Illinois: Office of the Superintendent of Instruction, 1972, unpublished manuscript.

Martin, E. Individualism and behaviorism as future trends in educating handicapped children. Exceptional Children, 38, 1972, 517-526.

McIntyre, M. "Emerging trends and issues in early childhood education". ERIC System: ED 103 111, 1975.

McLaughlin, M., Bussard, R., Smith, M., and Cohen, D. The Effects of Title I, ESEA: An Exploratory Study. Cambridge, Massachusetts: The Center for Educational Policy Research, 1971, (ERIC System: ED 073 216).

McLaughlin, D., Gilmartin, K., and Rossi, R. Controversies in the Evaluation of Compensatory Education: Summary of Recommendations. Palo Alto, Calif: American Institutes for Research, 1977.

McLaughlin, D., Gilmartin, K., and Rossi, R. Controversies in the Evaluation of Compensatory Education. Palo Alto, California: American Institutes for Research, 1977.

McLaughlin, M. "Parent Involvement in Compensatory Education Programs." Cambridge, Mass: The Center for Educational Policy Research (n.d.).

Messick, S. The Standard Problem: Meaning and Values in Measurement and Evaluation. American Psychologist, October, 1975, pp. 955-966.

Morrison, G. Early Childhood Education Today. U.S.A.: Bell and Howell, Co., 1976.

Mosback, E., with others. Analysis of Compensatory Education in Five School Districts: Final Report, Vol I: Summary; Vol. II: Case Studies. Washington, D. C.: General Electric Company - TEMPO, 1968.

Mussen, P. Couger, J., and Kagan, J. Child Development and Personality. New York: Harper and Row, 1969.

National Academy of Education. Improving Educational Achievement. Report of the National Academy of Education Committee on Testing and Basic Skills to the Assistant Secretary for Education, Washington, D. C: March 1, 1978.

National Center for Educational Statistics. Consolidated Program Information Report: Pupil Participation, Staffing and Expenditures in Federally Aided Programs Operated by School Districts, 1973. Washington, D. C.: U. S. Government Printing Office, 1976.

Neal, T. Organizing the states for children's services. Compact, 7, 1973, 8-10.

Nedler, S. Parent Education and Training: Literature Review. Austin, Texas: Southwest Educational Development Library, 1973.

Nuttal, E., and Gomes, L. "Early Identification of Children with Special Needs: A Survey of Current Screening Practices in Massachusetts Public Schools." Boston, MA: Boston College, 1975, unpublished ms.

Office of Planning, Budgeting and Evaluation. A Study of Compensatory Reading Programs, Technical Summary. Washington, D. C.: Elementary and Secondary Programs Division, 1977.

Ogura, J. Personal Communication, U. S. Office of Education, Title I Program Office, Washington, D. C., 1977.

Osborn, D. Early Childhood Education in Historical Perspective. Athens, Georgia: Educational Associates, 1975.

Parker, R., and Knitzer, J. Day Care and Preschool Services: Trends and Issues. Atlanta, Georgia: Avatar Press, 1972.

Parnell, D. Upgrading the downgrade. Compact. 7, 1963, 22-24.

Pelavin, S. H., and David, J. L. Evaluating Long-Term Achievements: An Analysis of Longitudinal Data from Compensatory Education Programs. California: Educationa Policy Research Center, Stanford Research Institute, 1977.

Piaget, J. The Psychology of Intelligence. New Jersey: Littlefield, Adams, 1966.

Piaget, J. The genetic approach to the psychology of thought. Journal of Educational Psychology, 52, 1961, 275-281

Piaget, J. Judgment and Reasoning in the Child. New York: Harcourt, Brace, 1928.

Piaget, J. The Language and Thought of the Child. New York: Harcourt, Brace, 1926.

Pierce, W. The Children's Needs Assessment Handbook. Colorado: The Education Commission of the States, 1976. (Rept. # 16)

Pierson, D. The Brookline Early Education Project: Model for a new educational priority. Childhood Education, 50, 1974, 132-135.

Pikunas, J. Human Development: An Emergent Science. New York: McGraw Hill, 1976.

Planar Corporation. The Silken Purse: Legislative Recommendations for Title I of the Elementary and Secondary Act. Submitted to DHEW/ASPE, 1973.

Raizen, S. and Bobrow, S. Designs for a National Evaluation of Social Competence in Head Start Children. Santa Monica, Calif: Rand, 1974.

Rampton, C. A commitment in jeopardy? Compact, 7, 1973, 3-4.

Riles, W. California's approach: the first seven years count the most. Compact, 7, 1973, 12-14.

Rosen, P., Horne, E. Tests for Spanish-Speaking Children, An Annotated Bibliography. New Jersey: Educational Testing Service, Head Start Collection, 1971.

Rossi, R. McLaughlin, D., Campbell, E., and Everett, B. Summaries of Major Title I Evaluations, 1966-1976. Palo Alto, Calif: American Institutes for Research, 1977.

Ryan, S. (ed.) A Report on Longitudinal Evaluations of Preschool Programs. Volume I: Longitudinal Evaluations. DHEW Publication No. (OHD) 74-24, 1974.

Salasin, S. (ed.) Evaluation. 1-4, 1972-1977, Minneapolis, Minn: Program Evaluation Resource Center, National Institute of Mental Health.

Schlossman, S. Before Home Start: Notes toward a history of parent education in America, 1897-1929. Harvard Educational Review, 46, 3, 1976, 436-467.

Schlossman, S. "The Parent Education Game: The Politics of Child Psychology in the 1970's". Teachers College Record, May, 1978, forthcoming.

Seng, M. (ed.) The Education Index, XVI, 1965-1966, N.Y.: H.W. Wilson, Co., 1966.

Seng, M. (ed.) The Education Index XV . 1965, N.Y.: H.W. Wilson, Co. 1965

Seng, M. (ed.) The Education Index, XII. 1959-1961, N.Y.: H.W. Wilson Co., 1961

Sigel, I. Where is preschool education going: Or are we without a road map? In Educational Testing Service (eds.) Assessment in a Pluralistic Society, Princeton, N.J.: Educational Testing Service, 1972, 99-116.

Simon, A., and Boyer, E. (eds.) Mirrors and Behavior II: An Anthology of Observation Instruments. Philadelphia, Pa: Research for Better Schools, 1967-1970 .

Snyder, L., Apolloni, T., and Cooke, T. Integrated settings at the early childhood levels: The tools of non-retarded peers. Exceptional Children, 43, 1977, 262-270.

Smith, M., and Bissell, J. Report Analysis: the impact of Head Start. Harvard Educational Review, 1970, 40, 51-104.

Spodek, B. Judging the maturity of early childhood education: the child is the father to the person. Contemporary Education, 45, 1974, 249-253.

Stearns, M. Evaluation of the Field Test of Project Information Packages. Volume I Summary Report. California: SRI, Int'l, 1977.

Stearns, M. Early Education: still in its infancy. American Education, 6, 1970, 3

Stebbins, L., et al. Education as Experimentation: A Planned Variation Model. Cambridge, MA: Abt Associates, April, 1977.

Stebbins, L., Proper, E., and Ames, N. Problems in Measuring Program Implementation Paper presented at the Annual Meeting of the American Research Association, San Francisco, California: April, 1976.

Stenner, A. J., Hunter, E. L. Bland, J. D. and Cooper, M. L. "The Standardized Growth Expectation: Implications for Educational Evaluation". Paper presented at the annual meeting of the American Educational Research Association, Toronto, Canada, 1978.

Stocker, J. Early Childhood Education: Current Trends in School Policies and Programs. Arlington, Virginia: National School Public Relations Association, 19

Stone, A. "American Status of Early Childhood Education, School Year 1974-1975." ERIC System, ED 118 272, 1976.

Sulzbacher, S., and Kenowitz, L. At last, a definition of learning disabilities we can live with? Journal of Learning Disabilities, 10, 1977, 67-69.

Tallmadge, G., and Wood, C. User's Guide: ESEA Title I Evaluation and Reporting System. California: RMC Research Corporation, October, 1976, Draft.

Terman, L. M., & Merrill, M. A. Manual for the Third Revision of the Stanford Binet Intelligence Scale (Form L-M). Boston, MA: Houghton-Mifflin, 1972.

Thomas, T., and Pelavin, S. Patterns in ESEA Title I Reading Achievement. Stanford Research Institute, Educational Policy Center, California: March, 1976

Thomas, T., and Pelavin, S. Patterns in Title I Reading Achievement. California: Stanford Research Institute, 1976.

Tizzard, J., Moss, P., and Perry, J. All Our Children: Preschool Services in a Changing Society. London: Temple Smith New Society, 1976.

Towner, I. The Education Index, I. 1929-1932. NY: The H. W. Wilson, Co., 1932.

Trismen, D., Waller, M., Wilder, G. A Descriptive and Analytic Study of Compensatory Reading Programs, Final Report, Volume I. Princeton, New Jersey: Educational Testing Service, 1975.

Trismen, D., Waller, M., Wilder, G. A Descriptive and Analytic Study of Compensatory Reading Programs, Appendices A, B, C. Princeton, New Jersey: Educational Testing Service, Final Report, Volume I, 1975.

Trismen, D., Waller, M., and Wilder, G. A Descriptive and Analytic Study of Compensatory Reading Programs - Final Report Volume II. Princeton, N.J.: Educational Testing SErvice, 1976.

Trismen, D., Waller, M., Butler-Nalin, K. A Descriptive and Analytic Study of Compensatory Reading Programs, Addendum I to the Final Report, Volume I. Princeton, N.J.: Educational Testing Service, 1976.

Trismen, D., Waller, M. A Descriptive and Analytic Study of Compensatory Reading Programs, Addendum II to the Final Report, Volume I. Princeton, N.J.: Educational Testing Service, 1976.

U. S. Bureau of the Census. School Enrollment - Social and Economic Characteristics of Students: October, 1977. (Series P-20, No. 321), Washington, D. C: USGPO, 1978.

U. S. Department of Health, Education and Welfare - Office of Education. A Practical Guide to Measuring Project Impact on Student Achievement. Number 1 in a Series of Monographs on Evaluation in Education, Washington, D. C.: U. S. Government Printing Office, 1975.

U. S. Department of Health, Education and WElfare, Office of Education. A Procedural Guide for Validating Achievement Gains in Educational Projects. Number 2 in a Series of Monographs on Evaluation in Education. Washington, D. C.: U. S. Government Printing Office, 1976.

United States Office of Education. Educational Programs that Work. San Francisco, Cal: Far West Laboratory for Educational Research and Development, 1976.

United States Office of Education. "Request for a Proposal: Development of Evaluation Models for Title I Programs in Early Childhood Education, 77-79". Washington, D. C.: 1977.

Villaume, J., and Haney, W. The Follow-Through Planned Variation Experiment, Volume V Appendix: Analysis of Interim Follow Through Evaluation Reports. Cambridge, Mass: Huron Institute, 1977.

Walker, D. Socioemotional Measures for Preschool and Kindergarten Children. San Francisco: Calif: Jossey-Bass, 1973.

Wargo, M., Tallmadge, G., Michaels, D., Lipe, D., and Morris, S. ESEA Title I A Re-Analysis and Synthesis of Evaluation Data from Fiscal Year 1965 Through 1970. AIR-27400-3/72-FR. Final Report. American Institutes for Research, California: March, 1972.

Washington Research Project/NAACP Legal Defense and Education Fund. Title I of ESEA: Is it Helping Poor Children? Washington, D. C.: Washington Research Project, 1969.

Weber, E. Early Childhood Education: Perspectives on Change. Worthington, Ohio: Charles A. Jones, 1970.

Weber, E. The function of early childhood education. Young Children, 28, 1973, 265-275.

Weber, E., and Gerhardt, L. Early childhood education issues and answers. Theory into Practice, 12, 1973, 69-70

Weinberger, P. Trends versus fads. Teacher, 94, 1977, 91-95.

Wenar, C. Personality Development From Infancy to Adulthood. N.Y.: Houghton-Mifflin, 1971.

Werner, H. Comparative Psychology of Mental Development. International Universities Press, 1957.

Wisler, C., Burns, G., and Iwamato, D. "Follow Through Redux: A Response to the Critique by House, Glass, McLean, and Walker." Harvard Educational Review, 48:2, May 1978, 171-185.

Wolfensberger, W. The Principle of Normalization in Human Services. Toronto: National Institute on Mental Retardation, 1972.

Zeitlin, S. Kindergarten Screening: Early Identification of Potential High Risk Learners. Springfield, Illinois: C. Thomas, Publisher, 1976.

Zigler, E. Myths and facts: a guide for policy makers. Compact, 7, 1973, 18-21.